

MECHANISMS WITHIN CONTEXTS: FIRST STEPS TOWARDS AN INTEGRATIVE APPROACH

Diego Zilio¹

Federal University of Espírito Santo (UFES)

ABSTRACT: In a series of discussions that later became known as the debate between mechanism and contextualism, some behavior analysts vehemently criticized the mechanistic proposition in favor of the contextualist alternative by arguing that the latter would be more consistent with radical behaviorism. The purpose of this article is to revisit this debate, but this time from the perspective of explanatory mechanism. Explanatory mechanism has been gaining attention in several domains of science, from molecular biology to the social sciences, to the point of being considered the most influential position in contemporary philosophy of science. In addition, considering that the debate between contextualism and behaviorism predates the explanatory mechanism revival in philosophy of science, it is important to explore the place of behavior analysis in this new mechanism framework. For this purpose, I will start with the presentation of explanatory mechanism by taking into account three points: the definition of mechanism, the conception of mechanistic explanation, and the methodological strategies of mechanistic research. After that, I will argue that explanatory mechanism not only is immune to the criticisms made in contextualism literature, but also that it is possible (and probably desirable) to adopt explanatory mechanism within contextualism.

Keywords: Behavior Analysis; Radical Behaviorism; New Mechanism; Contextualism; Explanation.

Mechanisms within contexts: First steps towards an integrative approach

In a series of texts and discussions that later came known as the “debate between mechanism and contextualism” (Morris, 1997), some behavior analysts vehemently criticized the mechanistic proposition in favor of the contextualist alternative. They argued that contextualism would be more consistent with radical behaviorism (Hayes & Hayes, 1992; Hayes, Hayes & Reese, 1988;

¹ Diego Zilio was funded by The Tutorial Education Program (PET) from the Ministry of Education (MEC). Correspondence about this paper should be addressed to Diego Zilio, Department of Social Psychology and Development, Federal University of Espírito Santo (UFES). Graduation Program in Psychology (PPGP – UFES). Avenida Fernando Ferrari, 514. CEMUNI VI. Goiabeiras, Vitória – ES. Brazil CEP: 29075-910. E-mail: dzilioufes@gmail.com

ZILIO

Morris, 1988, 1992, 1993a, 1993b, 1997; Reese, 1993). The purpose of this article is to revisit this debate, but now from the perspective of explanatory mechanism (I took the term from Bunge, 2004, but I use it here to refer to the main characteristics presented in contemporary mechanism literature).

Unlike Morris (1993), whose interest was “in the meaning of mechanism for those who call behavior mechanistic analysis, not in the meaning of mechanism for those who call themselves mechanists” (p. 256), my goal is to explore the meaning of mechanism for those who define themselves “new mechanists,” (Wimsatt, 2018) and then evaluate whether the conception of mechanism found is compatible with radical behaviorism as well as immune to the criticism of mechanism found in the contextualist literature. Mechanistic thinking has been gaining attention and advocates in several domains of science. The explanatory mechanism literature focuses mainly on biology as a model science to discuss mechanism, especially neuroscience (e.g., Bechtel, 2008, 2009; Bechtel and Richardson, 2010; Bickle, 2003a, 2003b, 2007, 2008a; Silva, 2007). Neuroscience will also be the exemplar case discussed here. However, explanatory mechanism can also be found in sociology (Bunge, 1997, 2004), particularly in the branch called “analytical sociology” (Hedström, 2005; Hedström & Bearman, 2009) or “social scientific mechanism” (Glennan & Illari, 2018a). A recently published handbook on mechanisms and mechanical philosophy (Glennan & Illari, 2018b) presents examples of mechanistic explanations in physics, molecular, evolutionary and systems biology, biomedicine, ecology, neuroscience, cognitive science, psychology, sociology, history, economics, computer science, and engineering. The journal *Theory & Psychology* published in 2019 a special issue on mechanisms in psychology (Milkoski, Hohol, & Nowakowski, 2019a) with the title “Mechanisms in psychology: The road towards unity?” Introducing the issue, Milkoski, Hohol, and Nowakowski (2019b) start by saying that “in contemporary philosophy of science, the mechanistic framework ... is currently the most influential approach to explanation in life, behavioral, cognitive, and social sciences” (pp. 567-568). In this context, one can at least assume that mechanism seems to be a trend topic in philosophy of science, from biology to sociology. Considering in addition that the literature on explanatory mechanism was published after the debate between contextualism and behaviorism, which took place in the earlier 1990s, I think it is important to explore the place of behavior analysis in this new mechanism. For this purpose, our trajectory begins with the presentation of explanatory mechanism by taking into account three points: the definition of mechanism, the conception of mechanistic explanation, and the methodological strategies of mechanistic research. In the following, I will argue that explanatory mechanism not only is immune to the criticisms made in contextualism literature but also that it is possible (and probably desirable) to adopt explanatory mechanism within contextualism.

Defining mechanism

Machamer, Darder, and Craver (2000) presented one of the first definitions of mechanism by using examples of contemporary biological sciences: “Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination

conditions” (p. 3). Mechanisms are formed by “entities” which, in turn, have functions (“activities”). These entities are “organized” in such a way as to produce regular changes from an “initial” point, prior to the changes, to a post-change “final” point. As it stands, the definition of Machamer, Darden, and Craver (2000) generates more questions than answers, because it presents concepts (e.g., entity, activity, organization) that also need clarification. In fact, the authors devote much of the text to the treatment of these key concepts in the definition of mechanism, but before starting this discussion, it is necessary here to present other definitions of “mechanism.” The aim is to show that there is some agreement regarding the definition of the concept.

Bechtel and Abrahamsen (2005) present a definition very similar to the previous one: “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). The authors also highlight the fact that mechanisms have a “function,” that is, they are responsible for the occurrence of a “phenomenon” – Machamer, Darden, and Craver (2000) use the term “changes,” and this function occurs because of the “organization” of the “components” (the “entities”) of the mechanism and their “operations” (the “activities”).

In turn, Bechtel and Richardson (2010) mention other characteristics: “A machine is a composite of interrelated parts, each performing its own functions, that are combined in such a way that each contributes to producing a behavior of the system” (p. 17). First, the authors associate “mechanism” with the idea of “system.” The mechanism consists of components that have their own functions and are organized in such a way to produce the behavior of the mechanism (or “system”) itself. The second feature highlighted by Bechtel and Richardson (2010) is the “interrelationship” between the components of the mechanism and their functions. The occurrence of the function of the mechanism (i.e., the occurrence of the phenomenon) depends on the interaction of the constituent components. It is only in this interaction that we can assess what would be the “activity” or “function” of a given component in that system (Machamer, Darden & Craver, 2000).

Craver presents definitions that also encompass the aforementioned characteristics: “Mechanisms are collections of entities and activities organized together to do something” (2005, p. 385); and “mechanisms are entities and activities organized such that they exhibit the explanandum phenomenon” (2007, p. 6).

ZILIO

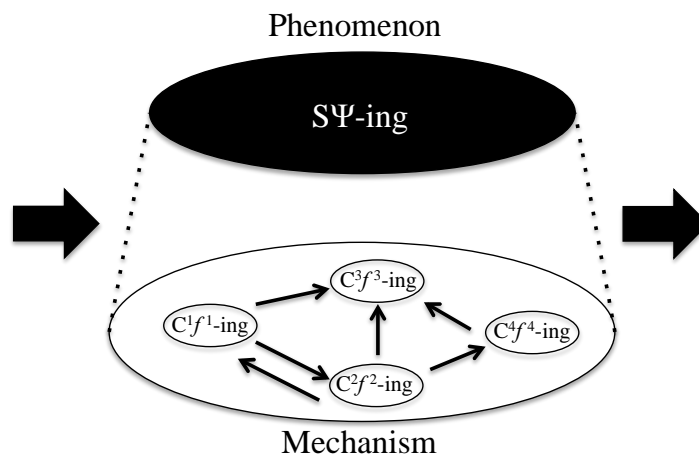


Figure 1. *The phenomenon and its mechanism.* Adapted from Craver (2007).

At the top of *Figure 1*, adapted from Craver (2007), we have the phenomenon to be explained. Specifically, the behavior (Ψ) of the system (S). The “-ing” suffix indicates action in progress. Therefore, the phenomenon is a process. Large arrows outside the mechanism also suggest this non-static character. The notion of “state” of the mechanism as something finished and inert has little contribution in this perspective. Macherson, Darden, and Craver (2000) state: “Often, mechanisms are continuous processes that may be treated for convenience as a series of discrete stages or steps” (p. 12). At the base of the figure, we find the components of the mechanism ($C_1, C_2, C_3, C_4 \dots, C_n$) and their functions or activities ($f_1, f_2, f_3, f_4, \dots, f_n$). The arrows indicate the interaction between the components and the organization of the system. Another important characteristic evidenced by the figure is that the mechanism is not only responsible for the occurrence of the phenomenon, but also constitutes it. The phenomenon is not something outside the mechanism. On the contrary, it is composed by the components of the mechanism. An essential characteristic of the mechanistic metatheory is that the explanation of the phenomenon (i.e., the behavior of the system) results largely from the analysis of the system itself (of its components, activities and organization) and not from entities and/or activities external to the system, although the importance of the “context” in which the system is inserted is recognized (Bechtel, 2008, 2009).

From these considerations, it is possible to conclude that there are four elements in the definition of mechanism. First, there are the “entities,” or “components,” which constitute it. Again, with Craver (2007):

Entities are the components or parts in mechanisms. They have properties that allow them to engage in a variety of activities. They typically have locations, sizes, structures, and orientations. They are the kinds of things that have masses, carry charges, and transmit momentum. They also act in a variety of

ways, by binding to other objects, opening and closing, and diffusing. (pp. 5-6)

Bechtel (2008) describes the entities as being "... structural components of a mechanism" (p. 14). In sum, mechanisms are made up of components. As we will see later, much of the challenge of mechanistic research is to decompose the mechanism by locating the components that constitute it and are responsible for the occurrence of the phenomenon. In addition, these components have "activities," or "functions." This is the second element we should take into account in the definition of mechanism. In the words of Craver (2007):

Activities are the causal components in mechanisms. I use the term "activity"... merely as a filler term for productive behaviors (such as opening), causal interactions (such as attracting), omissions (as occurs in cases of inhibition), preventions (such as blocking), and so on. In saying that activities are productive, I mean that they are not mere correlations, that they are not mere temporal sequences, and, most fundamentally, that they can potentially be exploited for the purposes of manipulation and control. (p. 6)

Activities are the functions performed by the components in the context of a specific mechanism. Consequently, they are not intrinsic to the components: "To see an activity as a function is to see it as a component in some mechanism, that is, to see it in a context that is taken to be important, vital, or otherwise significant" (Machamer, Darden, & Craver, 2000, p. 6). The functions are "causal," in the sense of having a fundamental role in the occurrence of the phenomenon. Craver (2007) seems to maintain a manipulationist positioning on causality, according to which "... variable X is causally relevant to variable Y in conditions W if some ideal intervention on X in conditions W changes the value of Y (or the probability distribution over possible values of Y)" (p. 94). That is, the role of the components and their functions in a given mechanism is evaluated from experimental manipulations. If manipulation of a component "X" (e.g., making it impossible for it to perform its function) results in effects on the behavior of the system "S" (e.g., the non-occurrence of the phenomenon), then "X" is likely to be relevant to "S," which makes it part of the mechanism. In addition, as we are dealing with relations of constitution – the mechanism is not only responsible for the occurrence of the phenomenon, but also constitutes it – the manipulation of the behavior of the system "S" as a whole can also cause changes in the activity of the component "X."

The possibility of manipulation is essential in the search for the elements of the mechanism. Here, we find the criterion of demarcation of the mechanism, which Craver (2005) calls "mutual manipulation." If the manipulation of a component affects the behavior of the system and the manipulation of the system affects the activity of the component, then we can assume that the component is part of the system.

To clarify this question, Craver (2005) distinguishes the "components" of the mechanism from the "background" elements. For example, the digestive system acts in the "background" when the focus is the strengthening of synaptic connections in the cortex. Digestion is essential for the survival of the organism and, by extension, for the occurrence of strengthening of synaptic connections, but it is not part of the mechanism responsible for the strengthening of synaptic

ZILIO

connections. The demarcation is established from manipulation criteria. The manipulation of components of the digestive system does not directly affect the strengthening of synaptic connections. Although it may have effects on the behavior of the organism and specifically on the digestive process, the artificial manipulation of saliva secretion does not directly increase or decrease the strengthening of synaptic connections in the cortex. The opposite is also true, hence the description of the criterion as “mutual manipulation.”

However, Craver (2005) agrees that demarking a mechanism is in the end a pragmatic endeavor: “No doubt, the distinction between background conditions and components is often drawn on pragmatic grounds. However, such pragmatic decisions can be made on an objective base” (p. 157). The criterion of mutual manipulation is not intended to exhaust all possible scenarios of reciprocal manipulation and influence between distinct components and mechanisms. It is only a relatively safe objective basis from which we can draw the boundaries of the mechanisms and locate their components.

The third element we should take into account in the definition of mechanism is the “organization” and “interaction” of its components and their respective functions. Craver (2007) affirms that “... the entities and activities in mechanisms are organized together spatially, temporally, causally, and hierarchically... The behavior of the mechanism as a whole requires the organization of its components” (p. 6). The components, their functions, and their interactions have to be spatially and temporally organized in a specific way so that the phenomenon is produced. The organization of these components is not mere arbitrary aggregation.

Finally, the fourth element to be considered in the definition of mechanism is the “phenomenon,” or “behavior,” of the system itself. Craver and Bechtel (2006) argue that this would be the “phenomenal” aspect of mechanistic research: every mechanism is associated with a phenomenon. This point is very important, because it is not possible to delimit a “mechanism” without first defining the specific “phenomenon” that one intends to study. Mechanisms are always mechanisms of *something*, and it is this “something” that directs the parameters of the initial and terminal boundaries of the mechanism. In the words of Craver (2007): “The boundaries of the mechanism are fixed by reference to the phenomenon to be explained” (2005, p. 390); and “the description of mechanisms is ineliminably perspectival” (2007, p. 259), which means that “functional description is ineliminably perspectival in the sense that it relies ultimately on decisions by an observer about what matters or is of interest in the system they study” (Craver, 2013, p. 133). It should be noted, therefore, that the search for mechanisms has to start with a definition of the phenomena object of study and this has mainly to do with scientist own interest of study. Without this definition, there are no criteria for establishing the limits of the mechanism itself. This also means that there seems not to be mechanisms as natural-kind entities. On the contrary, mechanisms are analytical products from scientists’ behavior of decomposing nature in the process of studying it.

A case example: neuroscience mechanisms

The characteristics of explanatory mechanism as presented by Bechtel (2008, 2009), Bickle (2003a, 2003b, 2007, 2008a), Craver (2007), Silva (2007),

among others, have neuroscience as the main source of analysis. Bickle (2008a) described his strategy as follows:

... to turn straight to the published experimental reports, here to neuroscience's primary experimental literature, with an eye to first doing purely descriptive metascience – to making explicit particular features of the science that typically remain implicit in the practices themselves, and burdened with as few prior metaphysical or normative epistemological convictions as we can be. The result of such an investigation would be a description of what the scientists are actually doing. (p. 15)

Bickle's metascientific strategy is to begin with as little normative or metaphysical assertions about scientific practice as possible, and then to proceed with the study of the scientist's behavior. Bickle (2003a, 2003b, 2007, 2008a) openly renounces the "arm-chair" philosophy of science, concerned with delineating normative criteria of scientific practice without consulting (or minimally consulting) the experimental contingencies themselves. His alternative consists of "... examining detailed scientific case studies in an effort to understand the nature of a relationship (like reduction) in actual scientific practice, rather than imposing philosophical intuitions on an analysis of what that relation 'has to be'" (Bickle, 2007, p. 277). Bickle's metascience is in accordance with the "naturalistic" proposal of Bechtel (2008):

... the naturalist proposes that we should examine how scientific inquiry is conducted by actual scientists and in doing so avail ourselves of the resources of science. That is, the philosopher of science would focus on securing data about how scientists work and developing theoretical accounts that are tested against that data. Although such an approach cannot independently specify norms for doing science, it can draw upon scientists' own identification of cases that constitute good and bad scientific practice and use these to evaluate theories about how science works, as well as to evaluate work within the sciences that are the objects of study. (p. 7)

The pragmatic element of Bickle's and Bechtel's proposals is also found in Silva (2007) and Craver (2007). Silva (2007), a neuroscientist, proposes a research program called "science of science," whose objective would be "... development of pragmatic, validated general principles for increasing the efficiency of science, just as the ultimate goal in medical research is the understanding and systematic development of pragmatic practices that improve the efficiency of prevention and treatments" (p. 204). Craver (2007), in turn, maintains that the practical criteria of intervention and control should be adopted in explanatory mechanism. In his words: "One way to justify the norms that I discuss is by assessing the extent to which those norms produce explanations that are potentially useful for intervention and control" (2007, pp. ix-x). Not only should we observe what the scientist has been doing in order to find the theoretical-philosophical foundations behind his research strategies, but we should also assess the pertinence of those strategies and fundamentals having their practical consequences as a criterion. Craver (2007), Silva (2007), Bickle (2003a, 2003b, 2007, 2008a), and Bechtel (2008) have adopted this perspective and, even differing in specific points in their ideas, have come to the unanimous conclusion that explanatory mechanism best suits neuroscientific practice.

ZILIO

In what way, then, explanatory mechanism is related to neuroscience? Craver (2002) gives us the answer:

Mechanisms, as they are understood in contemporary neuroscience, are collections of entities and activities organized in the production of regular changes from start or setup conditions to finish or termination conditions.... The entities in neuroscience include things like neurons, neurotransmitters, brain regions, and mice. The activities are the various doings in which these entities engage: neurons fire, neurotransmitters bind to receptors, brain regions process, and mice navigate mazes. Activities are the things that entities do; they are the productive components of a mechanism, and they constitute the stages of mechanisms. When neuroscientists speak generally about activities, they use a variety of terms; activities are often called "processes," "functions," and "interactions." When they speak specifically about activities, they use verbs and verb forms; they speak of attracting and repelling, phosphorylating and hydrolyzing, binding and breaking, and firing and releasing. The entities and activities composing mechanisms are organized; they are organized such that they do something, carry out some task or process, exercise some faculty, perform some function or produce some end product. I will refer to this activity or behavior of the mechanism as a whole as the role to be explained by the description of the mechanism. (p. 84)

A good example of explanatory mechanism in neuroscience is in the description of "long-term potentiation" (LTP). It is a relevant phenomenon in the study of neurophysiological mechanisms related to learning and memory (Morris, 2003). The discovery of LTP strengthened the hypothesis that memory, as "neurophysiological retention," would be related to modifications in synaptic efficacy (Zilio, 2013a). It is known that these modifications occur in several areas of the nervous system, such as the frontal cortices, amygdala, and hippocampus (Martin, Grimwood & Morris, 2000). There are several types of LTP (Martin, Grimwood & Morris, 2000; Teyler & DiScenna, 1987). Some involve modifications in presynaptic neurons, such as increasing the amount of neurotransmitters released to each action potential and the morphological development of more axons, thus increasing the number of synaptic connections with dendrites of postsynaptic neurons (Kandel, 2001; Squire & Kandel, 1999); others, in turn, involve modifications in the postsynaptic neurons, such as the increase of receptor channels of neurotransmitters and the morphological development of more dendrites, thus increasing the number of synaptic connections with axons of the presynaptic neurons (Kandel, 2001; Squire & Kandel, 1999). Thus, when we speak of "LTP" in the most general sense, we are referring to an abstraction based on several distinct neurophysiological mechanisms, but which, however, have a similar function: the modification of synaptic efficacy.

Let's take as an example the process of change of synaptic efficacy between the presynaptic neurons constituted by the *Schaffer Collateral fibers* and the postsynaptic neurons located in the *CA1 area* of the hippocampus (Rosenzweig, Breedlove, & Watson, 2005). The initial condition of the mechanism has, among other things, the *NMDA* channel blocked by Mg^{2+} molecules in the postsynaptic neuron. In this phase, only *AMPA* receptor acts as

a channel for excitation of the neuron through the influx of Na^+ , which, in turn, occurred thanks to the excitatory neurotransmitter glutamate released by the presynaptic neuron. The repeated activation of *AMPA* receptor central produces the depolarization of the dendritic membrane of the postsynaptic neuron. Depolarization causes the Mg^{2+} molecules to shift and, as a consequence, *NMDA* channel is unblocked. With the channel unblocked, the glutamate released by the presynaptic axon allows the infusion of Ca^{2+} ions through *NMDA* channel, initiating the LTP process. Ca^{2+} ions, combined with *calmodulin* (*CaM*) molecules, activate several second messengers, among them the protein *CaM Kinase* (*CaMK*), responsible for increasing the sensitivity of *AMPA* receptor to glutamate by means of its phosphorylation, besides causing movement to the membrane of other *AMPA* receptors that until then were disabled. The increase of *AMPA* channels implies more glutamate receptors, which means that there will be greater influx of Na^+ . Finally, the greater influx of Na^+ will be responsible for the increase in the excitatory potential of the postsynaptic neuron (Rosenzweig, Breedlove & Watson, 2005).

In addition to increasing *AMPA* channels, the influx of Ca^{2+} through *NMDA* channel has another consequence related to synaptic efficacy. Activation of protein *CaM Kinase* (*CaMK*), protein Kinase C (*PKC*), and protein Tyrosine Kinase (*TK*), through the influx of Ca^{2+} , eventually activates retrograde messengers that send signals to the presynaptic neuron. These retrograde messengers do not have vesicles with specific substances waiting to be released, since the synaptic connections are unidirectional (Squire & Kandel, 1999). Retrograde messengers are substances synthesized from the activation of protein kinases. In the specific case of LTP described here, the retrograde messenger is nitric oxide (*NO*) gas (Boehning & Snyder, 2003). The postsynaptic neuron diffusely releases *NO* out of the membrane, and its reception by the presynaptic neuron results in the facilitation of the release of the neurotransmitter glutamate. But there is one condition: this facilitation only occurs if the presynaptic neuron is activated during *NO* reception (Squire & Kandel, 1999). Thus, even diffuse, the release of *NO* ends up affecting only the presynaptic neurons that are part of LTP.

The above synaptic enhancement process is usually related to short-term potentiation (Frey, Huang & Kandel, 1993; Frey & Morris, 1998). Late long-term potentiation (L-LTP or “late-LTP”), in turn, is associated with structural changes in neurons dependent on protein synthesis (Frey, Huang & Kandel, 1993; Frey & Morris, 1998; Kandel, 2001; Squire & Kandel, 1999). At this point, the neuromodulators start acting, especially dopamine. When dopamine is released into the synaptic cleft, its respective receptors on the postsynaptic neurons are activated, event which is responsible for the activation of the adenylyl cyclase enzyme. Activation of adenylyl cyclase, in turn, results in the production of a second messenger, *cyclic adenosine monophosphate* (*cAMP*), by synthesizing the adenosine triphosphate (*ATP*) molecule. Increased *cAMP* concentration in the postsynaptic neuron triggers several biochemical events, including phosphorylation of the *cAMP-dependent protein kinase* (*PKA*). *PKA* phosphorylation results in its translocation into the nucleus of the neuron, where it phosphorylates the *cAMP response element binding protein* (*CREB*). *CREB* is a transcription modulating protein responsible for regulating neuron growth (Silva *et al.*, 1998). In this case, *CREB* phosphorylation eventually triggers a

ZILIO

genetic activation sequence whose final result is the structural change of the postsynaptic neuron. Specifically, there is growth of new dendrites that will form synaptic connections with the axons of presynaptic neurons that have been active throughout the L-LTP process.

Usually, in neuroscience articles and books those process descriptions are followed by illustrative figures of the mechanism being described. This makes the description (and the mechanism) a lot easier to comprehend (Craver, 2007). Instead of mathematical formulations, or general laws, mechanistic explanations provide mechanisms descriptions (Wright & Bechtel, 2007) and the figures present us with the components (or entities) and their functions, as well as illustrating the organization of the mechanism as a whole. Even though not using figures in my description of LTP mechanism, it is possible to find the defining elements of the mechanism in this brief example. First, there is the delimitation of the phenomenon object of analysis. We are talking about the mechanisms associated with “modification of synaptic efficacy.” This is the phenomenon. The definition of the object of study, which may change among observers/scientists according to his or her interested of study (Craver, 2013), establishes the initial and final boundaries of the mechanism. Another relevant point is that the mechanistic explanation is constitutive, which means that the mechanism is responsible for the occurrence of the phenomenon and also constitutes it. It is not an external cause producing the phenomenon. All steps described above are responsible for the change in synaptic efficacy, but, at the same time, they are also the process of synaptic efficacy change itself.

We also find in the example the “entities,” or “components,” of the mechanism, as well as their functions. Mg^{2+} , Na^{+} , NO, CaM, CaMK, AMPA, NMDA, ATP, cAMP, PKA, CREB, among others, are all components of the mechanism, each one having particular functions. Frequently, the function is already implicit in the very description of the process: molecules “block,” “activate,” and “excite”; channels are “open”; membranes are “depolarized” because of the activation of “receptors”; proteins “phosphorylate,” are “synthesized,” and “modulate” the “transcription” of genes; their “expression,” in turn, “regulates” the “growth” of neurons; all of these activities jointly promote and constitute the phenomenon described as the “increase in synaptic efficacy.”

It is important to emphasize that the functions of the components exist only in the context of the mechanism (Machamer, Darden, & Craver, 2000). That is, they are not necessarily intrinsic to the components. It is only possible to know the function of a component from the analysis of the relation between the components themselves. In the case of the “increase in synaptic efficacy” mechanism, the release of dopamine activates the adenylyl cyclase enzyme and thus triggers the L-LTP process. It is also known that the release of dopamine is associated with the presentation of reinforcing stimuli. Neurons of the ventral tegment area (VTA), for example, are known to diffusely project axons into various brain areas, including the motor associative cortex, the motor nucleus sub-cortical (nucleus accumbens), and the areas of the frontal cortex that have axons derived from sensory and sensory-associative cortices, and for releasing dopamine into these areas when their dendrites are stimulated. The occurrence of reinforcing stimuli is responsible for activating VTA neurons (e.g., Donahoe & Palmer, 1994; Guerra & Silva, 2010; Phillmore, 2008; Schultz, 1998, 2006;

Zilio, 2013a). Because of these factors, it is assumed that there is a relation between the occurrence of reinforcing stimuli, dopamine release, and the increase in synaptic efficacy. This would be one of the ways in which learning and retention of new behavioral relations would occur (Zilio, 2013a). Nevertheless, this is not the only function of dopamine. In the context of the mechanism related to changes in synaptic efficacy, dopamine may have this neuromodulation function; it may even be associated with reinforcement; but in other mechanisms, that is, in other “contexts,” dopamine may acquire distinct functions. For example, it plays a central role in the regulation of renal functions and prolactin secretion, physiological processes not necessarily related to learning, memory, and reinforcement (Zilio, 2013a). In this way, it is wrong to assign intrinsic functions to the components of a mechanism. It is only in the whole, that is, in the “context” of the mechanism, that we can locate the function of its components.

Finally, the description of the process of increase in synaptic efficacy also takes into account the spatial and temporal organization of the components and their respective functions. The organization of the components of the mechanism is not arbitrary. Here, the maxim is that the whole is not just the sum of its parts (Craver, 2007). It is necessary to take into account the spatial and temporal constraints of the mechanism to understand how the phenomenon is produced. For example, we described that the postsynaptic neuron diffusively releases NO out of the membrane and that the reception of NO by the presynaptic neuron results in the facilitation of the release of the neurotransmitter glutamate. However, this facilitation only occurs if the presynaptic neuron is activated during NO reception (Squire & Kandel, 1999). In addition, in the case of L-LTP, the new postsynaptic dendrites only form connections with the axons of the presynaptic neurons that were active throughout the process (Squire & Kandel, 1999). In this context, we say that LTP involves “associative” synapses, since it is dependent on the coactive functioning of pre- and postsynaptic neurons (Martin, Grimwood & Morris, 2000; Rosenzweig, Breedlove & Watson, 2005). These factors can be considered as organizational characteristics of the mechanism. That is, for the phenomenon to be produced, these conditions have to be satisfied.

Mechanistic explanation

Once clarified the defining characteristics of mechanism, the next step is to discuss the very conception of explanation in explanatory mechanism. I will approach this issue maintaining neuroscience as the exemplar case. According to Bechtel and Abrahamsen (2005), the definition of mechanism already reveals the purpose of the neuroscientific explanation: “The very conception of a mechanism lays out the tasks involved: the scientist must identify the working parts of the mechanism, determine what operations they perform, and figure out how they are organized so as to generate the phenomenon” (p. 432). Craver (2005) proposes a quite similar description:

The goal in describing a mechanism . . . is to describe this productive flow, without gaps from the beginning of the mechanism to its end. When we show how a mechanism works, we describe the relevant parts of the mechanism

ZILIO

and their relevant properties, and we show how their activities are organized together so that the working of the mechanism is transparent and unmysterious. As such, the search for mechanisms is driven by the goals of replacing black boxes . . . and filler terms (e.g., filler action verbs such as causes, activates, inhibits, modulates) with descriptions of mechanisms. (p. 386)

Therefore, to explain is to provide a description of the mechanism from its initial to final condition – it is to show how the phenomenon is produced. But it is not a description of a fixed “state” or “structure.” The mere description of the organs of the digestive system does not capture the mechanism of digestion; a description of the presynaptic and postsynaptic neurons does not comprise the mechanism related to the increase in synaptic efficacy. The mechanistic description should highlight the “productive flow,” that is, how the components of the mechanism, each with its respective function and organized in a given way, produce the phenomenon object of analysis.

Craver (2005) also states that the mechanistic explanation seeks to eliminate the “black boxes” and the “filler” terms of the descriptions. “Filler” terms are terms commonly used to describe the functions of the components of the mechanism. In the case of increase in synaptic efficacy, for example, proteins have been said to “phosphorylate,” and genes are “expressed” and “transcribed.” However, these terms only “fill” the description of the mechanism, but do not explain it. On the contrary, it is necessary to explain the process of “phosphorylation,” “transcription,” and “gene expression.” For Craven (2007), these explanations would refer to more details about the mechanism.

The “black boxes” are the gaps between the starting and the final condition of the mechanism. Explaining a mechanism is analogous to solving a puzzle. We often put together the peripheral pieces of the image without knowing what its center looks like. The same goes for the explanation of a mechanism. At one point in the description of the L-LTP process, for example, we had “put together” the pieces related to dopamine release into the synaptic cleft, adenylyl cyclase enzyme activation, cAMP production, and PKA phosphorylation. The final result would still be the structural change of the postsynaptic neuron. But there is a “black box,” that is, a gap in the process between PKA phosphorylation and the structural change of the neuron. This gap is filled by the description of other components of the mechanism and their respective activities: PKA translocation into the nucleus of the postsynaptic neuron, CREB phosphorylation and, finally, transcription of neuronal growth regulating genes.

However, it is paramount to emphasize that this explanation without “gaps” is an idealized goal. It is not possible to cover all components, all activities, and all organizational aspects of the mechanism; that is, it is not possible to eliminate all “black boxes” and replace all the “filler” terms. As stated by Craver (2006): “Few if any mechanistic models provide ideally complete description of a mechanism. In fact, such descriptions would include so many potential factors that they would be unwieldy for the purposes of prediction and control and utterly unilluminating to human beings” (p. 360). Mechanisms are models created by scientists through their interactions with the world, especially with the phenomenon object of study. As they pretend to explain the phenomenon by describing the mechanism that constitutes it and

which is, therefore, responsible for its occurrence, the models also provide subsidies for effective manipulation of the phenomenon it represents.

In this context, Craver (2006, 2007) distinguishes between “possible” (how-possible) and “effective” (how-actually) models. Possible models do not have a necessary and direct relation to the phenomenon they are trying to explain. In the words of the author (2006): “They describe how a set of parts and activities might be organized such that they exhibit the explanandum phenomenon. One can have no idea if the conjectured parts exist and, if they do, whether they can engage in the activities attributed to them in the model” (p. 361). The cognitive models and constructs criticized by Skinner (Zilio, 2016b) and endorsed by the cognitivist metatheory (Baars, 1986; Baars & Gage 2010) can be included in this category. In turn, the “effective” models are composed by elements directly extracted from empirical neurophysiological research. Once again with Craver (2006): “How-actually models describe real components, activities, and organizational features of the mechanism that in fact produces the phenomenon. They show how a mechanism works, not merely how it might work” (p. 361).

For Craver (2006), effective models can be classified in a continuum between sketches and ideally complete descriptions of the mechanism. The sketches consist of models providing incomplete descriptions of the mechanism regarding some component, activity, or organizational aspect (Robins & Craver, 2009). They usually have “black boxes” and “filler terms” (Craver, 2005, 2006, 2007). At the end of the continuum there is the complete description of the mechanism. However, as stated earlier, this description is impractical and, because of its complexity, could not provide conditions for prediction and control of the phenomenon. Because of this, in the words of Robins and Craver (2009), “... models frequently drop details that are irrelevant in the conditions under which the model is to be used” (p. 59). Our brief description of the mechanism related to the increase in synaptic efficacy is, by no means, complete. However, it is possible to manipulate the phenomenon efficiently by taking into account the characteristics described. We could, for example, artificially introduce dopamine into the synaptic cleft, triggering the neuronal growth process in L-LTP (cf. Wise, 2004); or eliminate the α and Δ forms of CREB through gene manipulation, resulting in the impediment of L-LTP due to the lack of this component essential to the process (cf. Silva *et. al.*, 1998).

Since a complete and exhaustive description of the mechanism is impracticable, the decision of what will be put in and left out of the description have manipulation as the primary guide. A model is sufficiently “complete” if it promotes conditions for effective manipulation of the phenomenon. Again, with Craver (2007):

According to the manipulationist account, explanatory texts describe relationships between variables that can be exploited to produce, prevent, or alter the explanandum phenomenon. Merely being able to manipulate a phenomenon, of course, is not sufficient to explain it. People made babies long before they understood how DNA works. But the wider the range of possible manipulations, and the deeper one’s knowledge of how such manipulations change the explanandum phenomenon, the more complete is the explanation. (pp. 100-101)

ZILIO

Bickle (2006) seems to defend a similar position: “Experiments only claim a successful explanation, or a successful search for a cellular or molecular mechanism, ... when they successfully intervene at the lower level and then measure a statistically significant behavioral difference under precise controlled conditions” (p. 137). The author is more specific in his description, since it emphasizes cellular and molecular manipulations and the evaluation of the effects of these manipulations on the behavior of the subject. For example, behavioral contingencies whose establishment depends on hippocampal activity, such as the task in Morris’s aquatic labyrinth, do not cause long-term changes in the behavior of rats genetically manipulated to be deficient in the production of α and Δ forms of CREB in the hippocampus (Silva *et al.*, 1998). It is said, for example, that the subject’s ability to “learn” behavior related to spatial “memory” was compromised by this genetically manipulated deficiency (cf. Silva *et al.*, 1998). This is an example of how manipulation at the molecular level can affect the behavior of the organism as a whole. However, for Bickle (2006), this is part of the “explanation” of the behavior since its possible to establish a functional relation by means of control between these variables.

Although it is an essential criterion in the mechanistic explanation, the possibility of manipulation is not an explanation in itself. It is not enough to produce the phenomenon (Craver, 2007). It is necessary to know the components, their respective functions, and organization of the mechanism. It is not enough to genetically manipulate the rat and see the effect on behavior. It is necessary to understand why such manipulation has this effect on behavior. To do so, we need to know how the mechanism works. This knowledge, in turn, enables new means of control, and this must also be taken into account in the classification of the “completeness” of the explanatory model (Craver, 2007). A model related to the increase in synaptic efficacy that suggests that the manipulation of dopamine concentration may influence the occurrence of L-LTP is less complete than a model that, in addition to informing this possibility, also presents the possibility of CREB manipulation.

Craver (2007) synthesizes the relation between manipulation and explanation in explanatory mechanism:

One need not be able to derive the phenomenon from a description of the mechanism... one needs to know how the phenomenon changes under a variety of interventions into the parts and how the parts change when one intervenes to change the phenomenon. When one possesses explanations of this sort, one is in a position to make predictions about how the system will behave under a variety of conditions. Furthermore when one possesses explanations of this sort, one knows how to intervene into the mechanism in order to produce regular changes in the phenomenon. Explanations in neuroscience are motivated fundamentally by the desire to bring the CNS under our control. (p. 160)

Just as it is not possible to deduce the existence of “water” as liquid substance with specific properties from the formula “H₂O,” the mechanistic research does not aim to derive the phenomenon from the description of its mechanism. The goal is to describe its components, including their respective functions and organizational properties. Prediction and control of phenomena is the main criteria to support (or validate) a mechanistic explanation.

In sum, mechanistic explanations consist in the description of the mechanism from its starting to its final condition. It should show how the components of the mechanism, each with its respective function and organized in a given way, produce the phenomenon object of analysis. However, this description is never complete. It is not possible to exhaust all that is possible to know about the mechanism. There will always be some “black box” or some “filler” term. Therefore, we speak of “models” of mechanisms. In the words of Bechtel (2008): “what the scientist advances is a representation of a mechanism – a construal of it – not the mechanism itself” (p. 18). The manipulation criterion has to be taken into account in the classification of mechanistic models. Models are valid when they allow effective manipulation of the phenomenon they intend to explain. A model is more “useful” or more “complete” in comparison to others when enables more and better control conditions.

Observational research

According to Silva (2007; see also Bickle 2006, 2008b, 2009; Silva & Bickle 2009), neuroscientific research can be classified in at least four kinds. First, there is purely observational research that does not seek to intervene directly in the phenomenon of study: “*Observation* refers to experiments designed to determine whether one natural phenomenon regularly follows another. Observation experiments do not intentionally alter the phenomena” (Silva & Bickle, 2009, p. 101). Thus, for example, several studies indicate that there is correlation between changes in synaptic plasticity in several brain areas and learning of specific behaviors (cf. Silva, 2007). Brain imaging studies, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), fall into the correlational or, to use Silva’s (2007) term, observational category. Usually, the experimental subject is submitted to a given behavioral contingency (the “cognitive task”) at the same time as the neural activities are measured through these devices.

Finding this correlation is an important starting point for formulating new research questions. Attesting the correlation, the next step is to describe the mechanism behind it. These correlations also function as support in the validation of hypotheses already based on information about the mechanism. For example, the amygdala seems to be part of the neurophysiological mechanism associated with fear conditioning (LeDoux, 2000). The research that supports this hypothesis usually involves some kind of manipulation (e.g., lesion, genetic manipulation, and substance infusion) of amygdala in order to see the effects on behavior. Nonetheless, when the subject is a human being, this approach is not advisable, or, better saying, ethically acceptable. In this case, noninvasive observational methods are more indicated. For instance, it is known that when human beings are subjected to fear conditioning, there is increase in amygdala activity, and subjects with bilateral amygdala lesions are not as susceptible to fear conditioning as non-injured subjects (Büchel & Dolan, 2000; Labar *et. al.*, 1998; Phelps, 2006).

In short, correlations do not reveal the characteristics of the mechanism responsible for the phenomenon. Bickle (2008b) says that “meeting the Observation Principle does not by itself establish that the hypothesized mechanism is part of the causal nexus generating the behavioral measures” (p.

ZILIO

43). However, they indicate a possible path to follow in the search for this mechanism, besides justifying the generalization of hypotheses developed from research with other animals for the case of the human beings.

Positive and negative manipulations

The following two principles or aspects of neuroscientific research involve the manipulation of elements supposedly belonging to the mechanism and the analysis of the effects of such manipulation. The manipulation can be negative or positive (Bickle, 2006, 2008b, 2008c, 2009; Silva, 2007; Silva & Bickle 2009). Negative manipulation encompasses the experiments in which "... the probability of one natural phenomenon is decreased and the effect on another (typically on its probability) is tested" (Silva & Bickle, 2009, p. 104). We have already briefly mentioned the example of the genetically-modified rats with deficiency in the production of α and Δ forms of CREB and the effects of this long-term learning disability on tasks that depend on hippocampal activity (Silva *et al.*, 1998). In this case, the manipulation is negative because it decreases the probability of the presence of CREB. The effect is also negative in this case, but this is not the rule for all negative manipulations. Depending on which the outcome is, we say that the manipulated element has either a stimulatory (i.e., its presence increases the probability of occurrence of the phenomenon) or an inhibitory (i.e., its presence reduces the probability of occurrence of the phenomenon) function (Craver, 2007).

Positive manipulations, in turn, encompass the experiments "... wherein the probability of one natural phenomenon is increased and the effect on another (typically on its probability) is tested" (Silva & Bickle, 2009, p. 106). Another example already presented here is artificially introducing dopamine into the synaptic cleft between the neurons of the Schaffer Collateral fibers and the neurons of the CA1 area of the hippocampus, a manipulation that increases the probability of occurrence of neuronal growth in L-LTP (Wise, 2004). The manipulation is "positive," since there is the addition of something. The effect here is positive (i.e., increased probability of occurrence of the phenomenon) and, therefore, we say that dopamine would have a "stimulatory" function in the production of the phenomenon.

Integration

Finally, the fourth and final principle of neuroscientific research is "integration," defined as "... a general class of scientific activity that does not directly involve either manipulation or observation, but focuses instead on hypothesizing, ordering, and cross-referencing connections between phenomena" (Silva & Bickle, 2009, p. 108). Bickle (2009) describes it as the activity of "... relating different levels of theory and explanation" (p. 412). In other words, it is the integration of information about a given phenomenon arising from different areas and/or levels of analysis. For example, information about behavioral contingencies, neural circuits and systems, and cellular and molecular characteristics of the brain are all essential to understanding the mechanisms behind the production of behavior (Zilio, 2016a).

Explanatory mechanism and contextualism criticisms

At first glance, the association between explanatory mechanism and radical behaviorism may seem impossible, if we assume the latter being more equivalent to a contextualist worldview. However, one of the central questions, if not the main one, in the debate between contextualism and mechanism concerns the very definition of the “mechanism” under analysis (cf. Morris, 1993a). Usually, the term is given a simplistic and pejorative meaning: man would be a “machine,” such as a clock or a mousetrap, composed of discernible parts independent of each other, but which, when organized, produce a given function. This definition can be found in Hayes and Hayes (1992):

The root metaphor of mechanism is the machine. A machine (such as a lever) consists of discrete parts (e.g., a fulcrum and lever), a relation among these parts (e.g., the lever must sit atop the fulcrum), and forces to make the parts operate (e.g., pressing down on one end of the lever produces a precisely predictable force at the other end). (p. 229)

In fact, there are similarities between this definition of mechanism and the one presented in explanatory mechanism. Both argue that mechanisms are made up of components; that each component has a function; and that, when organized in a certain way, they are responsible for the production of a phenomenon. The difference is in the details. Critics attribute a number of characteristics to mechanism that are not present in explanatory mechanism.

Causality and explanation

First, mechanism would be associated with a conception of causality incompatible with radical behaviorism, according to which causal relations between events would be linear and unidirectional, occurring through the transmission by contact of some kind of force between the links of the causal chain. There would be no causal gaps: all elements of the chain are connected to each other (Carrara, 2002, 2004; Chiesa, 1992, 1994, Hayes, 1988; Morris, 1988, 1992, 1993a; Moxley, 1992; Overton & Reese, 1973; Reese, 1993; Reese & Overton, 1970). Chiesa (1994), for example, affirms that mechanism “... implies linear causation, traceable backward to a single causal event, as if all causal relations once set in motion proceeded in a straightforward one-to-one fashion” (p. 108), and that “... the chain metaphor requires that, like a machine, causal relations be contiguous in both space and time, that gaps between cause and effect be filled by a sequence of events standing in a relation of succession” (p. 116). Morris (1988), in turn, places this definition in the context of stimulus-response psychology: “As for causation, the elements are said to act on one another as do physical forces, the results of which are chain-like connections between, or sequences of, stimuli and responses” (pp. 298-299). In another moment, the author (1993a) highlights the linearity and unidirectionality of the conception: “In adhering to unidirectional and linear causality, the mechanist's explanatory model is an asymmetric, one-way relation between independently defined causes and effects” (p. 34).

Despite the problems that may arise from these characteristics, it is sufficient to say that they are not present in explanatory mechanism. As seen

ZILIO

earlier, explanatory mechanism adopts a manipulationist conception of causality. A component of a mechanism is “causal” in the sense of being a relevant piece for the occurrence of the phenomenon. Relevance, in turn, is attributed from experimental manipulations. Craver (2007) describes this conception as follows:

... causal relevance, explanation, and control are intimately connected with one another. This is particularly true in biomedical sciences, such as neuroscience, that are driven not merely by intellectual curiosity about the structure of the world, but more fundamentally by the desire (and the funding) to cure diseases, to better the human condition, and to make marketable products. The search for causes and explanations is important in part because it provides an understanding of where, and sometimes how, to intervene and change the world for good or for ill. This connection between causation, explanation, and control is also reflected in the procedures that neuroscientists use to test explanations. These tests involve not only revealing correlations among the states of different parts of a mechanism but, further, intervening in the mechanism and showing that one has the ability to change its behavior predictably. (p. 93)

The scientific endeavor in explanatory mechanism is not about unraveling the nature of causal relations; it is about understanding the conditions under which scientists speak of “causal relations” and it is when they manipulate elements of the mechanism and see what follows. It is precisely in this context that the manipulationist conception of causality makes sense (Woodward, 2003). Craver (2007) emphasizes the connection between causality, explanation, and control. Explanations must show how the components of the mechanism, each with its respective function and organized in a given way, produce the phenomenon object of analysis. These components are not “causal” in the sense described in the aforementioned criticisms made in the behavior analysis literature. We say that a component has a “causal” role if its manipulation affects in some way the production of the phenomenon as a whole, but the relation between component and phenomenon does not have to be necessarily linear, contiguous, and unidirectional. Claiming that CREB phosphorylation “causes” the growth of new dendrites is equivalent to saying that CREB “manipulation” had effects on the growth of new dendrites.

For Skinner (1957), to explain is to discover the variables relevant to the occurrence of a phenomenon and the way in which they relate. In this case, explanation would not be mere description or narrative of events. Explaining demands manipulation of variables under analysis. This is the essence of the notion of explanation as a “description” of functional relations between events (Marr, 2003). One behavior analyst may say that there is a functional relation between the increase in the rate of responses belonging to a given class and the contingent presentation of a reinforcing consequence. One neuroscientist may say that there is a functional relation between CREB negative manipulation (e.g., eliminating α and Δ forms through gene manipulation) and the non-occurrence of L-LTP. In both cases, scientists manipulated independent variables and evaluated the effects of such manipulation on dependent variables.

In sum, according to this manipulationist conception of causality, linearity, unidirectionality and transmission of some kind of force by contact are not necessary conditions in order to classify a relation between events as

“causal.” The use of the term “cause” is associated with the manipulation of variables. In addition, the explanation of the phenomenon is the description of how the mechanism works (its components, activities, and organization). It is not a linear explanation in which the first link in the chain is the initial cause of the phenomenon to be found in the final link of the causal chain. The mechanism that explains the phenomenon is the phenomenon itself, analyzed and classified in components, functions, and organizational aspects. Bechtel and Richardson (2010) classify the non-linear mechanisms as “integrated” mechanisms:

Although one might prefer explanations in which the component tasks can be thought of as following a linear, sequential order, so that the contributions of each component can be examined separately, natural systems are not always organized in such a manner. Component tasks are often dependent on one another, so we cannot understand the operation of the system by imposing a linear order on it... In integrated systems, the explanation of the behavior of the whole system depends in a nonlinear way on the activities of the components and on the modes of interaction found within the system. (p. 202)

Gaps

However, what can be said about the inadmissibility of gaps in mechanistic explanations? These seem to be acceptable in behavior analysis explanations, therefore distancing it from explanatory mechanism. This question lead to the idea of “distance,” “mnemic,” or “historical” causality, usually associated with behavior analysis (cf. Marr, 1993, 2008; Marr & Zilio, 2013; Morris, 1988; Schaal, 2005). In general, mechanism would not be compatible with behavioral analysis because it accepts the existence of spatial and temporal gaps between functionally dependent events (Chiesa, 1994). The specific contingency “X” to which the subject was submitted at a given point in the past would be responsible for the actual occurrence of the behavior “Y.” There is a spatial and temporal gap between “X” and “Y.” The problem, however, lies in the meaning attributed to the “gap.” First, it may indicate an intrinsic feature of the relationship between “X” and “Y.” In the manipulationist conception of causality, the transmission of some kind of force by contact is not a necessary condition to classify a relation between events as “causal.” So there is no problem in locating “causal” relations at temporal distance. On the other hand, the gap may only indicate that the intermediate events between “X” and “Y” are unknown. If there is a functional relationship between CREB phosphorylation and the growth of new dendrites – events that are temporally and spatially separated – it is because there are intermediate events responsible for linking these events. There are no actual “gaps” between such events, which is not the same as assuming a linear a-historical causal relation between them. However, there may be gaps in the analysis made by scientists. The gaps are the “black boxes” and the “filler” terms mentioned earlier. We might not know what happens between CREB phosphorylation and the growth of new dendrites, but this “gap” should be eliminated, as far as possible, by doing further experiments and manipulations. It should be not different in the case of behavior analysis. The “gaps” are in the analysis and not in the relationships between events. For

ZILIO

instance, there are gaps in behavior analysis explanations that, for Skinner, should be filled through neurophysiological research (Zilio, 2013b, 2016a).

Thus, using the “gap” argument as a justification against mechanism in behavior analysis is, at least, problematic. The first meaning of gap seems to be a metaphysical one, concerning the nature of the relation between events. Beyond metaphysics, however, one could argue against this notion of gap on pragmatic grounds. Accepting the existence of real temporal and spatial gaps between events may discourage scientists from searching for more pieces of the puzzle that may help us to get a more complete picture of how the phenomenon is produced. This actually happened in the early days of behavior analysis, when neuroscience was not taken into account in the explanation of behavior (Zilio, 2016a, 2016b). On the other hand, if it is accepted that the gap is a characteristic of the analysis due to its inherent incompleteness (i.e., it does not exhaust all the components, functions, and organizational aspects related to the production of the phenomena it aims to explain), then the argument loses its force completely, because the existence of “gaps” is a trivial and even intrinsic characteristic of science in general. There are gaps in behavior analysis, and the mechanism explanations are never complete as we saw earlier: there will always be “black boxes” and “filler” terms or, in a general sense, gaps. Marr (2008) synthesizes the question in the following excerpt:

... there never was a need by behavior analysts to invoke “action-at-a-temporal-distance.” Changes in neural systems and environment-behavior interactions are ever ongoing—continuous variables operating within immensely complex feedback systems. If I meet an old high school buddy whom I’ve not seen for many years, I wouldn’t label the obvious physical changes of aging as “action-at-a-distance,” temporal or otherwise—why would I use that expression to characterize behavior changes (and their associated neural processes) evident over time? (p. 166)

Essentialism or atomism

Explanatory mechanism seems to be immune to the criticisms made by behavior analysts about the supposed notion of mechanistic causality. In fact, the manipulationist perspective seems to be in line with Skinner’s (1953/1965) own proposal of searching for “functional relations” instead of causal ones. But criticisms of mechanism are not limited to the question of causality. Mechanism is also accused of maintaining an “atomistic” or “essentialist” position with regard to the components of the mechanisms (Carrara, 2002; Hayes, 1988; Hayes & Brownstein, 1986; Hayes & Hayes, 1992; Hayes, Hayes, & Reese, 1988; Morris, 1988, 1992, 1993a; Overton & Reese, 1973; Reese, 1993; Reese & Overton, 1970). Morris (1988) describes atomism as follows:

Complex action is ... an associative compounding of the basic elements and their interrelations, in which identical response elements and identical stimulus elements are taken, respectively, to have identical meanings or functions—such is the character of a machine. In both cases, the whole can always be reconstituted in terms of its parts because the parts are unchanging. (p. 300)

Hayes and Hayes (1992) present similar description:

If we wished to understand a machine, we would need to disassemble it and identify the parts, relations, and forces that constitute it and its operation. Note also that when the machine is disassembled, the parts remain unchanged despite their independence from the rest of the machine. In other words, a spark plug is a spark plug whether screwed into a cylinder or sitting on the kitchen table. (p. 229)

They are focusing on the mechanistic research strategy characterized by the decomposition of the phenomenon, that is, by dividing it in components with functions and organized in particular ways. However, the problem does not seem to be in the decomposition strategy itself, but rather in the assumption that the components have intrinsic functions, regardless of the mechanism of which they are part. This is what Morris (1988) seems to assume when he claims that elements with identical physical characteristics would have the same function or meaning. By addressing the question, the author rightly argues that in behavioral relations the function of responses and stimuli should not be derived from their physical properties. On the other hand, in mechanism, the functions would supposedly be inherent to the components. In this scenario, it would be possible to discover the function of a given component without evaluating the mechanism as a whole. Besides, the component would maintain the same function despite the mechanism of which it is part. To paraphrase Hayes and Hayes (1992): the spark plug will maintain its function no matter the car (i.e., the mechanism with a whole) of which it is part and even when it is not part, although momentarily, of any car. The function would be intrinsic to it.

However, the comparison between cars and biological systems is unsatisfactory. Explanatory mechanism does not assume that functions are intrinsic to the components of the mechanisms. The components have functions only in the context of the mechanism of which they are part. There is no atomism or essentialism about the parts. As Craver (2007) said:

The parts of neural mechanisms are in many cases not so visible, not so readily distinguished from their surroundings; in some cases, they are widely distributed and dynamically connected, defying any attempts to localize functions to particular parts. In that case, the machine analogy provides a misleadingly simplistic view of the mechanisms in nature. (p. 4)

This idea is far from the automobile caricature view of mechanism. The point is not the non-existence of atomism in some kind of mechanistic explanation, but that this commitment is not necessary in explanatory mechanisms.

Truth criterion and realism

Another common criticism directed to mechanism is related to its epistemological commitment. Mechanism would be bound to the criterion of truth by correspondence, whereas radical behaviorism would defend the pragmatic criterion of “effective action” (Hayes, 1988, 1993; Hayes & Brownstein, 1986; Hayes & Hayes, 1992; Hayes, Hayes, & Reese, 1988; Morris,

ZILIO

1988; Reese, 1993). Hayes (1993) describes the criterion of truth by correspondence as follows:

... its truth criterion [of mechanism] is an elaborated form of simple correspondence. Because the mechanist assumes the presence of a world organized *a priori* into events, relations, and forces, truth is to be found in the way ever more ambitious verbal formulae reveal the assumed organization of the world. (p. 12)

In another text, Hayes and Hayes (1992) affirm that "... the goal of mechanistic research is the development of a model of the machinery that is assumed to exist" (pp. 229-230). In other words, a mechanistic model is true as long as it corresponds to the actual mechanism existing in the world. Mechanistic research seeks to unravel such real mechanisms, that is, how the world actually works. There would be a directly relation between the validity of the model and the approximation of reality: the closer to how things really are, the truer is the model. This criticism seems to merge the epistemological problem about the validation criterion of an explanation into the ontological problem of the real nature of things. This feature is explicit in Morris (1988):

The truth criterion of these causal mechanisms is correspondence: Given that knowledge in mechanism is knowledge about the nature of a realist ontology, the truth of that knowledge is found in the correspondences across domains of the activity of the machine or in predictions between what is said about the machine (e.g. hypotheses) and how the machine operates (e.g. confirmations). (p. 299)

For the critics (Hayes, 1988; Hayes & Brownstein, 1986; Hayes & Hayes, 1992; Hayes, Hayes, & Reese, 1988; Moxley, 1987), ontological realism would strengthen, in the epistemic realm, the hypothetical-deductive method, specifically the proposition of purely conceptual mechanistic models (i.e., hypothetical constructs) in order to explain the behavioral processes from which they were inferred. These models would be evaluated based on their correspondence with the observed facts. According to Hayes and Hayes (1992): "If such a model is shown to correspond to a range of relevant observations (especially if it is predictively verified) then it is said to be true" (p. 230). I do not intend to enter into metaphysical discussions, since the defense or critique of realism would not change the fact that, in explanatory mechanism, the criterion to evaluate an explanation is pragmatic (Craver, 2007), and its research strategies are not necessarily related to hypothetical-deductive models of explanation (cf. Baars & Gage, 2010). In addition, Bickle's metascientific proposal, Bechtel's naturalistic approach, Silva's "science of science," and Craver's pragmatism have removed any a priori metaphysical commitment from explanatory mechanism. Bickle (2008a) describes his strategy as follows:

... to turn straight to the published experimental reports, here to neuroscience's primary experimental literature, with an eye to first doing purely descriptive metascience – to making explicit particular features of the science that typically remain implicit in the practices themselves, and burdened with as few prior metaphysical or normative epistemological

convictions as we can be. The result of such an investigation would be a description of what the scientists are actually doing. (p. 15)

Bickle's metascientific strategy is to begin with as little normative or metaphysical assertions about scientific practice as possible, and then to proceed with the study of the scientist's behavior. Bickle (2003a, 2003b, 2007, 2008a) openly renounces the "arm-chair" philosophy of science, concerned with delineating normative criteria of scientific practice without consulting (or minimally consulting) the experimental contingencies themselves. His alternative consists of "... examining detailed scientific case studies in an effort to understand the nature of a relationship (like reduction) in actual scientific practice, rather than imposing philosophical intuitions on an analysis of what that relation 'has to be'" (Bickle, 2007, p. 277).

Bickle's metascience is in accordance with the "naturalistic" proposal of Bechtel (2008):

... the naturalist proposes that we should examine how scientific inquiry is conducted by actual scientists and in doing so avail ourselves of the resources of science. That is, the philosopher of science would focus on securing data about how scientists work and developing theoretical accounts that are tested against that data. Although such an approach cannot independently specify norms for doing science, it can draw upon scientists' own identification of cases that constitute good and bad scientific practice and use these to evaluate theories about how science works, as well as to evaluate work within the sciences that are the objects of study. (p. 7)

What have neuroscientists been doing in the laboratory? What have they been calling "explanation"? What strategies have worked to achieve a research objective? These are the questions, among others, that should guide the proposition of the neuroscientific metatheory. The pragmatic content of Bickle's and Bechtel's proposals is also found in Silva (2007) and Craver (2007). Silva (2007), a neuroscientist, proposes a research program called "science of science," whose objective would be "... development of pragmatic, validated general principles for increasing the efficiency of science, just as the ultimate goal in medical research is the understanding and systematic development of pragmatic practices that improve the efficiency of prevention and treatments" (p. 204).

Craver (2007), in turn, maintains that the practical criteria of intervention and control should be adopted in the proposition of the neuroscientific metatheory. In his words: "One way to justify the norms that I discuss is by assessing the extent to which those norms produce explanations that are potentially useful for intervention and control" (2007, pp. ix-x). Not only should we observe what the scientists has been doing in order to find the theoretical-philosophical foundations behind his research strategies, but we should also assess the pertinence of those strategies and fundamentals having their practical consequences as a criterion. Craver (2007), Silva (2007), Bickle (2003a, 2003b, 2007, 2008a) and Bechtel (2008) have adopted this perspective and, even differing in specific points in their ideas, have come to the unanimous conclusion that the mechanistic metatheory is the one which best suits neuroscientific practice. Neuroscience aims to understand neurophysiological

ZILIO

mechanisms, and this objective has obvious practical consequences: “the discovery mechanisms provides scientists with new tools to diagnose diseases, to correct bodily malfunctions, to design pharmaceutical interventions, to revise psychiatric treatments, and to engineer strains of organisms” (Craver, 2007, pp. ix-x).

This way of approaching science is familiar to behavior analysts. When invited to write a chapter on scientific method, rather than enumerating a set of norms of scientific conduct, Skinner (1956) described his own history as a researcher, highlighting the events that, for him, were important for the development of his area of research. The author (1958) was critical of a priori norms so commonly presented in philosophy of science texts:

Certain people – among them psychologists who should know better – have claimed to be able to say how the scientific mind works. They have set up normative rules of scientific conduct. The first step for anyone interested in studying reinforcement is to challenge that claim. (p. 99)

Moreover, for Skinner (1966/1969a), the criterion of validation of scientific knowledge is practical: “Science is in large part a direct analysis of the reinforcing systems found in nature; it is concerned with facilitating the behavior which is reinforced by them” (p. 143). The purpose of science would be to create conditions for effective manipulation of the natural world (Smith, 1992, Zilio, 2010). In summary, the criterion of “effective action,” is not particular to behavior analysis, but is also present in explanatory mechanism.

Explanatory mechanism seems to be immune to the criticisms against mechanism presented in the debate between “mechanism and contextualism.” For this reason, the simplistic view of the “machine man,” indiscernible from a clock or car, criticized by behavior analysts, is not adequate to characterize explanatory mechanism. In addition to that, there is not any a priori metaphysical commitment in explanatory mechanism. It is not a worldview in Pepper’s (1942) sense. It is not an ontological position about how the world works. Robins and Craver (2009) are clear about that:

The word [mechanism] should not be understood as implying adherence to any strict metaphysical system. Clearly neural mechanisms are not generally understood as machines that work only through motion, attraction and repulsion, or the transmission of conserved quantities. Nor are they generally understood as heroic simple machines, or machines that work according to the principles of Newtonian mechanics, or strictly deterministic systems in which laws of nature allow only one output for any input. (p. 42)

Mechanisms within contexts

At least seven characteristics defines the explanatory mechanism research strategy:

1. Delimitation of the phenomenon;
2. Decomposition of the phenomenon into analytical units - the mechanism and its components;

3. Selection of variables (units or components) of interest;
4. Analysis of the function of the selected unit(s) or component(s) in the context of the mechanism of which it is part, through observational and manipulation protocols;
5. Analysis of the organizational characteristics of the mechanism as a whole, through observational and manipulation protocols;
6. Explanation, that is, the description of how the phenomenon is produced from the functional and organizational relation between the components of the mechanism;
7. Integrate the mechanism explanation into a broader view (big-picture) related to the phenomenon.

Described in this way, explanatory mechanism seems to be compatible with behavior analysis. First, behavior is defined as a continuous flow process. In Skinner's words: "it is a process, rather than a thing ... It is changing, fluid, and evanescent" (1953/1965, p. 15); "Behavior is very fluid; it isn't made up of lots of little responses packed together" (1968, pp. 20-21). But studying a processual phenomenon like behavior demands breaking it down into behavioral events (that is, a portion of the behavioral flux for time $t1$ to $t2$), and units of analysis. Skinner (1953/1965) agrees: "In this way behavior is broken into parts to facilitate analysis" (p. 93); and also:

A second kind of system, to which the term is intended to refer here, is clearly exemplified by the systems encountered in physical chemistry. Such a system consists of an aggregation of related variables, singled out for the sake of convenient investigation and description from all the various phenomena presented by a given subject matter. In the case of behavior, a system in this sense can be arrived at only through the kind of experimental analysis . . . in which the parts or aspects of behavior which undergo orderly changes are identified and their mutual relations established (1938/1966, pp. 434-435).

Hackenberg (1996) directly associates behavior analysis research strategies with the mechanistic methodology: "When we isolate controlling variables, are we not, in a sense, dissecting some part of the world into its constituent parts? ... This sounds like the reductive-analytic path of mechanism" (p. 300).

Let's take the three-term contingency as example. It takes this name for decomposing behavioral relations into three specific components. Its purely descriptive formulation would be something like: (a) the current context in which responses occur; (b) the responses themselves; and (c) the consequences that follow (cf. Skinner, 1969b). It is also possible to use illustrations (so common in mechanism explanations) with components, arrows, and particular organizations (see, for instance, Moxley, 1982, 1984; Morris, 1992).

ZILIO

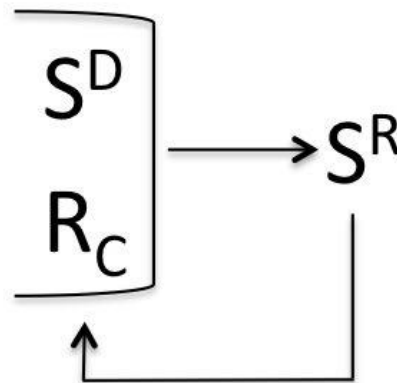


Figure 2. *The three-term contingency.*

Adapted from Marr (2013).

In *Figure 2*, adapted from Marr (2013), we have the components of the three-term contingency. The components' functions are in their description as “discriminative stimulus” (S_D), “operant class” (R_C), and “reinforcer” (S_R). In addition, the arrows and the bracket indicate the organizational characteristics. Specifically, the bracket signals that responses belonging to the same class occur in the presence of a given discriminative stimulus. The small horizontal arrow indicates that the occurrence of the response, in the presence of a given discriminative stimulus, leads to the production of a reinforcing stimulus. Finally, the arrow linking the S_R to the bracket indicates that the occurrence of the reinforcer modifies the relation between S_D and R_C . If S_R signals positive reinforcement, for example, we can have the effect of increasing the probability of occurrence of responses belonging to R_C in the presence of S_D . For Marr (2013), the three-term contingency can be viewed as a system of “interactive feedback,” since the consequence produced by a response on a given current context has a retroactive effect on the relation between responses and aspects of the current context in which it occurred. We could also classify it as an “integrated” system or mechanism, as defined by Bechtel and Richardson (2010): “In functionally integrated systems the behavior of the components is interdependent, so a change in the behavior of one part alters the behavior of others” (p. 150).

As argued above, integrated mechanisms are good examples of how functional (or causal) relations between components are not necessarily linear and unidirectional. *Figure 2* shows the incompatibility between the notion of linear and unidirectional causality and the dynamics of behavior selection. It also shows, on the other hand, the proximity between behavior analysis and explanatory mechanism. Proximity confirmed by Marr (2009): “Our field explores environment-behavior relations to develop principles accounting for the acquisition, maintenance, and dynamic changes in patterns of behavior—in other words, *mechanisms* of contingencies” (p. 112).

Of course, *Figure 2* is very simplistic. Not every component of the current context functions as discriminative stimuli, not every response is an

operant response, not every consequence is a reinforcer, and not every change in the world produced by a given behavioral class is a consequence to this class in a functional sense, although it might form new contexts that may engender new behaviors. In what he described as the unpacking of the three-term-contingency, Morris (1992) added a lot more components, functions, and organizational aspects in his illustration of the unpacked contingency (p. 15). Even though it was an effort to make the case for the contextualism in behavior analysis, it still is, in essence, an endeavor of explanatory mechanism as described in the seven characteristics presented earlier. Mattaini (1993, 1995, 1996) is also an informative and important example of effort to do the same but bringing social and cultural components to the mechanism. One could argue that the difference is only in degree or how much information (components, functions, and organizations) is inserted into the same mechanism. A simpler one, with only three components, can be very useful when dealing with the manipulation of particular variables in experimental settings. A broader one, which includes different components, functions, and organizations at different level of analysis (behavioral, social, cultural, economic, and so on) can be useful to promote an integrative approach. Either way, they still are mechanisms.

What seem to be driving the criticism of mechanism in contextualism literature are the assumed philosophical commitments of a particular kind of mechanistic world-view (e.g., essentialism, atomism, linear and a-historical causation, realism, truth by correspondence). As argued here, those commitments are not necessarily present in explanatory mechanism. In addition, the very attempt of bringing context to behavior analysis can be seen as an effort to expand the mechanism by adding new components, functions, and organizational aspects. In short, it seems possible to apply explanatory mechanism within contextualism. They are not incompatible at all. Actually, the systematicity of explanatory mechanism can bring some precision into contextualism in which “context” sometimes seems to be a filler term (almost a magical word) to refer to anything that is not behavior.

References

- Baars, B. J. (1986). *The cognitive revolution in psychology*. New York: The Guilford Press.
- Baars, B. J., & Gage, N. M. (2010). *Cognition, brain, and consciousness: Introduction to cognitive neuroscience* (2a. ed.). London: Academic Press.
- Barnes, D., & Roche, B. (1994). Mechanistic and contextualistic epistemology: a contradiction within behavior analysis. *The Behavior Analyst*, 17, 165-168.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York: Routledge.
- Bechtel, W. (2009). Looking, down, around, and up: Mechanistic explanation in psychology. *Philosophical Psychology*, 22(5), 543-564.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 421-441.

ZILIO

- Bechtel, W., & Richardson, R. C. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge: The MIT Press.
- Bickle, J. (2003a). Philosophy of mind and the neurosciences. In: S. Stich, & T. Warfield (Eds.), *The Blackwell guide to philosophy of mind* (pp. 322-351). Massachusetts: Blackwell Publishing.
- Bickle, J. (2003b). *Philosophy of neuroscience: a ruthlessly reductive account*. Dordrecht: Kluwer Academic Publishers.
- Bickle, J. (2006). Ruthless reductionism in recent neuroscience. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, 36(2), 134-140.
- Bickle, J. (2007). Who says you can't do a molecular biology of consciousness? In: M. Schouton, & H. L. De Jong (Eds.), *The matter of the mind: philosophical essays on psychology, neuroscience, and reduction* (pp. 275-297). Malden: Blackwell Publishing.
- Bickle, J. (2008a). Cognitive behaviors and molecular neurobiology: Explanations 'in a single bound'. In J. E. Burgos, & E. Ribes-Iñesta (Eds.), *The brain-behavior nexus: conceptual issues: Proceedings of the 10th biannual symposium on the science of behavior* (pp. 13-22). Guadalajara: Universidad de Guadalajara.
- Bickle, J. (2008b). Real reduction in real neuroscience: Metascience, not philosophy of science (and certainly not metaphysics!). In: J. Hohwy, & J. Kallestrup (Eds.), *Being reduced: new essays on reduction, explanation, and causation* (pp. 34-51). New York: Oxford University Press.
- Bickle, J. (2008c). The molecules of social recognition memory: Implications for social cognition, extended mind, and neuroethics. *Consciousness and Cognition*, 17, 468-474.
- Bickle, J. (2009). Cellular and subcellular neuroscience. In J. Symons, & P. Calvo (Eds.), *The Routledge companion to philosophy of psychology* (pp. 400-415). London: Routledge.
- Boehning, D., & Snyder, S. H. (2003). Novel neural modulators. *Annual Review of Neuroscience*, 26, 105-131.
- Büchel, C., & Dolan, R. (2000). Classical fear conditioning in functional neuroimaging. *Current Opinion in Neurobiology*, 10, 219-223.
- Bunge, M. (1997). Mechanism and explanation. *Philosophy of the Social Sciences*, 27(4), 410-465.
- Bunge, M. (2004). How does it work? The search for explanatory mechanisms. *Philosophy of the Social Sciences*, 34(2), 182-210.
- Carrara, K. (2002). *O mito da síntese experimental do comportamento: Reflexões a partir do behaviorismo radical e do contextualismo pepperiano*. Tese de Livre-docência. Marília: UNESP.
- Carrara, K. (2004). Causalidade, relações funcionais e contextualismo: Algumas indagações a partir do behaviorismo radical. *Interações*, 9(17), 29-54.
- Chiesa, M. (1992). Radical behaviorism and scientific frameworks: From mechanistic to relational accounts. *American Psychologist*, 47(11), 1287-1299.
- Chiesa, M. (1994). *Radical behaviorism: The philosophy and the science*. Boston: Authors Cooperative, Inc., Publishers.

- Craver, C. (2002). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, 69(3), 83-97.
- Craver, C. F. (2005). Beyond reduction: Mechanisms, multifield integration and the unity of neuroscience. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 373-395.
- Craver, C. (2006). When mechanistic models explain. *Synthese*, 153, 355-376.
- Craver, C. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.
- Craver, C. (2013). Functions and mechanisms: A perspectivalist view. In P. Huneman (Ed.), *Functions: Selection and mechanisms* (pp. 133-158). Netherlands: Springer.
- Craver, C., & Bechtel, W. (2006). Mechanism. In S. Sarkar, & J. Pfeifer (Eds.), *Philosophy of science: An encyclopedia* (pp. 469-478). New York: Routledge.
- Donahoe, J. (2013). Theory and behavior analysis. *The Behavior Analyst*, 36, 361-371.
- Donahoe, J. W., & Palmer, D. (1994). *Learning and complex behavior*. Boston: Allyn and Bacon.
- Frey, U., Huang, Y., & Kandel, E. (1993). Effects of cAMP simulate a late stage of LTP in hippocampal CA1 neurons. *Science*, 260, 1661-1664.
- Frey, U., & Morris, G. (1998). Synaptic tagging: Implications for late maintenance of hippocampal long-term potentiation. *Trends in Neuroscience*, 21(5), 181-188.
- Glennan, S., & Illari, P. (2018a). Introduction: Mechanisms and mechanical philosophies. In S. Glenna, & P. Illari (Ed.), *The routledge handbook of mechanisms and mechanical philosophy* (pp. 1-9). New York: Routledge.
- Glennan, S., & Illari, P. (2018b). *The routledge handbook of mechanisms and mechanical philosophy*. New York: Routledge.
- Guerra, L. G. C., & Silva, M. T. A. (2010). Learning processes and the neural analysis of conditioning. *Psychology & Neuroscience*, 3(2), 195-208.
- Hackenberg, T. D. (1996). When being a mechanist wasn't so bad: Reply to Moxley. *The Behavior Analyst*, 19(2), 299-300.
- Hayes, S. (1988). Contextualism and the next wave of behavioral psychology. *Behavior Analysis*, 23(1), 7-22.
- Hayes, S. (1993). Analytic goals and the varieties of scientific contextualism. In S. Hayes, L. Hayes, H. Reese, & T. Sarbin (Eds.), *Varieties of scientific contextualism* (pp. 11-27). Nevada: Context Press.
- Hayes, S. C., & Brownstein, A. (1986). Mentalism, behavior-behavior relations, and a behavior-analytic view of the purposes of science. *The Behavior Analyst*, 9(2), 175-190.
- Hayes, S., & Hayes, L. (1992). Some clinical implications of contextualistic behaviorism: the example of cognition. *Behavior Therapy*, 23, 225-249.
- Hayes, S., Hayes, L., & Reese, H. (1988). Finding the philosophical core: a review of Stephen C. Pepper's World hypotheses: a study in evidence. *Journal of the Experimental Analysis of Behavior*, 50(1), 97-111.
- Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology*. Cambridge: Cambridge University Press.

ZILIO

- Hedström, P., & Bearman, P. (2009). *The oxford handbook of analytical sociology*. New York: Oxford University Press.
- Kandel, E. (2001). The molecular biology of memory storage: A dialogue between genes and synapses. *Science*, *294*(5544), 1030-1038.
- LaBar, K. S., Gatenby, C., Gore, J., LeDoux, J., & Phelps, E. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron*, *20*, 937-945.
- LeDoux, J. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155-184.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*, 1-25.
- Marr, M. J. (1993). Contextualistic mechanism or mechanistic contextualism? The straw machine as tar baby. *The Behavior Analyst*, *16*(1), 59-65.
- Marr, M. J. (2003). The what, the how, and the why: the explanation of Ernst Mach. *Behavior and Philosophy*, *31*, 181-192.
- Marr, M. J. (2008). The abdication of belief: A comment on Gordon Foxall's replies to his critics. *Behavior and Philosophy*, *36*, 155-166.
- Marr, M. J. (2009). The natural selection: Behavior analysis as a natural science. *European Journal of Behavior Analysis*, *10*, 105-120.
- Marr, M. J. (2013). Tweedledum and tweedledee: Symmetry in behavior analysis. *Conductal: Revista Internacional de Interconductismo y Análisis de Conducta*, *1*, 16-25.
- Marr, M. J., & Zilio, D. (2013). No island entire of itself: reductionism and behavior analysis. *European Journal of Behavior Analysis*, *14*, 241-257.
- Martin, S., Grimwood, P., & Morris, G. (2000). Synaptic plasticity and memory: An evaluation of the hypothesis. *Annual Review of Neurosciences*, *23*, 649-711.
- Mattaini, M. (1993). *More than a thousand words: Graphics for clinical practice*. Washington: Nasw Press.
- Mattaini, M. (1995). Contingency diagrams as teaching tools. *The Behavior Analyst*, *18*(1), 93-98.
- Mattaini, M. (1996). Envisioning cultural practices. *The Behavior Analyst*, *19*(2), 257-272.
- Milkoski, M., Hohol, M., & Nowakowski, P. (2019a). Introduction. *Theory & Psychology*, *29*(5), 567-578.
- Milkoski, M., Hohol, M., & Nowakowski, P. (Ed.) (2019b). Mechanisms in psychology: The road towards unity? [Special issue]. *Theory & Psychology*, *29*(5).
- Morris, E. K. (1988). Contextualism: the world view of behavior analysis. *Journal of Experimental Child Psychology*, *46*, 289-323.
- Morris, E. K. (1992). The aim, progress, and evolution of behavior analysis. *The Behavior Analyst*, *15*, 3-29.
- Morris, E. K. (1993a). Behavior analysis and mechanism: on is not the other. *The Behavior Analyst*, *16*, 25-43.
- Morris, E. K. (1993b). Mechanism and contextualism in behavior analysis: just some observations. *The Behavior Analyst*, *16*(2), 255-268.
- Morris, E. K. (1997). Some reflections on contextualism, mechanism, and behavior analysis. *Psychological Record*, *47*(4), 529-542.

- Morris, R. (2003). Long-term potentiation and memory. *Philosophical Transactions: Biological Sciences*, 358(1432), 643-647.
- Moxley, R. (1982). Graphics for three-term contingencies. *The Behavior Analyst*, 5, 45-51.
- Moxley, R. (1984). Graphic discriminations for radical functional behaviorism. *Behaviorism*, 12, 81-95.
- Moxley, R. (1987). Three conceptual units for behavior. *The Behavior Analyst*, 10, 17-26.
- Moxley, R. (1992). From mechanistic to functional behaviorism. *American Psychologist*, 47(11), 1300-1311.
- Overton, W. F., & Reese, H. W. (1973). Models of development: Methodological implications. In J. R. Nesselrode, & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological issues* (pp. 65-86). New York: Academic Press.
- Palmer, D. C. (2004). Generic response classes and relational frame theory: Response to Hayes and Barnes-Holmes. *Journal of the Experimental Analysis of Behavior*, 82(2), 225-234.
- Pepper, S. C. (1942). *World hypotheses: A study in evidence*. Berkeley: University of California Press.
- Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, 57, 27-53.
- Phillmore, L. S. (2008). Discrimination: From behavior to brain. *Behavioural Processes*, 77, 285-297.
- Reese, H. W. (1993). Comments about Morris's paper. *The Behavior Analyst*, 16(1), 67-74.
- Reese, H. W., & Overton, W. F. (1970). Models of development and theories of development. In L. R. Goulet, & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory* (pp. 115-145). New York: Academic Press.
- Robins, S., & Craver, C. (2009). Biological clocks: Explaining with models of mechanisms. In J. Bickle (Ed.), *The Oxford handbook of philosophy and neuroscience* (pp. 41-67). New York: Oxford University Press.
- Rosenzweig, M., Breedlove, S., & Watson, N. (2005). *Biological psychology: An introduction to behavioral and cognitive neuroscience* (4a ed.). Massachusetts: Sinauer Associates, Inc.
- Schaal, D. W. (2005). Naming our concerns about neuroscience: A review of Bennett and Hacker's *Philosophical Foundations of Neuroscience*. *Journal of the Experimental Analysis of Behavior*, 84(3), 683-692.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80, 1-27.
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, 57, 87-115.
- Silva, A. (2007). The science of research: The principles underlying the discovery of cognitive and other biological mechanisms. *Journal of Physiology*, 101, 203-213.
- Silva, A., Kogan, J., Frankland, P., & Kida, S. (1998). CREB and memory. *Annual Review of Neurosciences*, 21, 127-148.
- Silva, A., & Bickle, J. (2009). The science of research and the search for molecular mechanisms of cognitive functions. In J. Bickle (Ed.), *The*

ZILIO

- Oxford handbook of philosophy and neuroscience* (pp. 92-126). New York: Oxford University Press.
- Skinner, B. F. (1956). A case history in scientific method. *The American Psychologist*, *11*(5), 221-233.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1958). Reinforcement today. *The American Psychologist*, *13*(3), 94-99.
- Skinner, B. F. (1965). *Science and human behavior*. New York: The Free Press. (Original work published 1953).
- Skinner, B. F. (1966). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts. (Original work published 1938).
- Skinner, B. F. (1968). [Interview]. In R. I. Evans, B. F. Skinner: *The man and his ideas*. New York: E. P. Dutton & CO., INC.
- Skinner, B. F. (1969a). An operant analysis of problem solving. In B. F. Skinner, *Contingencies of reinforcement: A theoretical analysis* (pp. 133-157). New York: Appleton-Century-Crofts. (Original work published 1966).
- Skinner, B. F. (1969b). The role of the environment. In B. F. Skinner, *Contingencies of reinforcement: A theoretical analysis* (pp. 3-22). New York: Appleton-Century-Crofts.
- Skinner, B. F. (1974). *About behaviorism*. New York: Alfred A. Knopf.
- Smith, L. D. (1992). On prediction and control: B. F. Skinner and the technological ideal of science. *American Psychologist*, *47*(2), 216-223.
- Squire, L., & Kandel, E. (1999). *Memory: From mind to molecules*. New York: Scientific American Library.
- Teyler, T., & DiScenna, P. (1987). Long-term potentiation. *Annual Review of Neurosciences*, *10*, 131-161.
- Wimsatt, W. (2018). Foreword. In S. Glenna, & P. Illari (Ed.), *The routledge handbook of mechanisms and mechanical philosophy* (pp. xiv-xvi). New York: Routledge.
- Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 1-12.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.
- Wright, C., & Bechtel, W. (2007). Mechanisms and psychological explanation. In P. Thagard (Ed.), *Handbook of the philosophy of science: Philosophy of psychology and cognitive science* (pp. 31-79). Amsterdam: Elsevier.
- Zilio, D. (2010). *A natureza comportamental da mente: behaviorismo radical e filosofia da mente*. São Paulo: Editora Cultura Acadêmica (Unesp).
- Zilio, D. (2013a). Behavioral unit of selection and the operant-respondent distinction: The role of neurophysiological events in controlling the verbal behavior of theorizing about behavior. *The Psychological Record*, *63*(4), 895-918.
- Zilio, D. (2013b). Filling the gaps: Skinner and the role of neuroscience in the explanation of behavior. *Behavior and Philosophy*, *40*, 33-59.

- Zilio, D. (2016a). On the autonomy of psychology from neuroscience: A case study of Skinner's radical behaviorism and behavior analysis. *Review of General Psychology, 20*(2), 155-170.
- Zilio, D. (2016b). Who, what and when: Skinner's critiques of neuroscience and his main targets. *The Behavior Analyst, 39*, 197-218.