NEO-MECHANISTIC EXPLANATORY INTEGRATION FOR COGNITIVE SCIENCE: THE PROBLEM OF REDUCTION REMAINS

INTEGRAÇÃO EXPLANATÓRIA NEOMECANICISTA PARA A CIÊNCIA COGNITIVA: O PROBLEMA DA REDUÇÃO PERMANECE

DIEGO AZEVEDO LEITE[1]
Università Degli Studi di Trento – Italy
diego.azevedo.leite@gmail.com

ABSTRACT: One of the central aims of the neo-mechanistic framework for the neural and cognitive sciences is to construct a pluralistic integration of scientific explanations, allowing for a weak explanatory autonomy of higher-level sciences, such as cognitive science. This integration involves understanding human cognition as information processing occurring in multi-level human neuro-cognitive mechanisms, explained by multi-level neuro-cognitive models. Strong explanatory neuro-cognitive reduction, however, poses a significant challenge to this pluralist ambition and the weak autonomy of cognitive science derived therefrom. Based on research in current molecular and cellular neuroscience, the framework holds that the best strategy for integrating human neuro-cognitive theories is through direct reductive explanations based on molecular and cellular neural processes. It is my aim to investigate whether the neo-mechanistic framework can meet the challenge. I argue that leading neo-mechanists offer some significant replies; however, they are not able yet to completely remove strong explanatory reductionism from their own framework.

KEYWORDS: Integration of cognitive science. Neuro-cognitive integration. Neo-mechanistic explanation. Explanatory pluralism. Explanatory integration.

RESUMO: *Uma das finalidades centrais da abordagem teórica neomecanicista para as ciências neural e cognitiva é a construção de uma integração pluralística de explicações científicas, permitindo uma autonomia explanatória fraca das ciências de mais alto nível, como a ciência cognitiva. Essa integração envolve a compreensão da cognição humana como processamento de informação ocorrendo em mecanismos neurocognitivos de múltiplos níveis, explicados por modelos neurocognitivos de múltiplos níveis. A redução explanatória neurocognitiva forte, no entanto, apresenta um desafio significativo para esta ambição pluralista e a autonomia fraca da ciência cognitiva dela derivada. Baseada em pesquisas na área atual da neurociência molecular e celular, essa abordagem teórica sustenta que a melhor estratégia para integrar teorias da neurocognição humana é através de explicações redutivas diretas baseadas em processos neurais moleculares e celulares. O meu objetivo é investigar se a estrutura teórica neomecanicista pode superar esse desafio. Eu argumento que os principais neomecanicistas oferecem algumas respostas significativas; porém, eles ainda não são capazes de remover completamente o reducionismo explanatório forte da sua própria estrutura teórica.*

PALAVRAS-CHAVE: *Integração da ciência cognitiva. Integração neurocognitiva. Explicação neomecanicista. Pluralismo explanatório. Integração explanatória.*

[1] PhD in Cognitive Science from the University of Trento, Italy (2018), in the line of research on Philosophy of Cognitive Science.

INTRODUCTION

The new *mechanistic theory of scientific explanation*, articulated in the end of the 20th century and the beginning of the 21st century, is one of the most important and influential contemporary theories of scientific explanation (BECHTEL; RICHARDSON, 1993/2010; BECHTEL; ABRAHAMSEN, 2005; MACHAMER; DARDEN; CRAVER, 2000; CRAVER; TABERY, 2015; GLENNAN; ILLARI, 2018). This theory is applied especially to the biological sciences, including cognitive science.

This neo-mechanistic framework applied to cognitive science is advocated by many contemporary influential authors (e.g. BECHTEL, 1994, 2007, 2008, 2009a, 2009b, 2009c, 2010, 2017; BECHTEL; WRIGHT, 2009; BOONE; PICCININI, 2016; CRAVER, 2002, 2007; KAPLAN, 2017; MILKOWSKI, 2016; PICCININI; BAHAR, 2013; PICCININI; CRAVER, 2011; PICCININI, 2007, 2012, 2015; THAGARD, 2006, 2009, 2018; WRIGHT; BECHTEL, 2007; ZEDNIK, 2018). Its central idea is that any human cognitive process is a kind of neural process, understood in terms of cognitive information processing and cognitive representation. Such processes can be decomposed and localized in brains as parts of a multilevel neural-biological mechanism. As a result, the framework suggests a path for a general 'pluralistic integration' of neuro-cognitive theories, allowing at the same time for some kind of weak autonomy[2] of scientific explanations in cognitive science.

However, central theoretical aspects of the neo-mechanistic framework for cognitive science remain highly controversial (cf. LEITE, 2018). Particularly, one of the major problems is about whether the neo-mechanistic framework can indeed provide the pluralist integration and weak explanatory autonomy of cognitive science it promises. One of the main obstacles comes from contemporary neuro-cognitive explanatory reductive approaches to the issue. John Bickle can be regarded as one of the most prominent representatives of contemporary neuroscientific explanatory reductionist ambitions (BICKLE, 2003, 2006, 2008, 2012, 2015, 2016). He contends that one should not see reduction in a negative manner in science, nor give up on the scientific general reductionist project, since reductionism can improve scientific disciplines mainly through theoretical unification and elimination of explanatory redundancy (BICKLE, 2003). He argues

[2] Bechtel has an article, published in 2007, titled "Reducing psychology while maintaining its autonomy via mechanistic explanations". In this work, he writes "I will argue in subsequent sections that the reductions achieved through mechanistic explanations are in fact compatible with a robust sense of autonomy for psychology" (2007, p. 174). Moreover, in his famous work of 2008, Bechtel writes: "Traditionally, arguments for the autonomy of psychology have appealed to the claimed multiple realizability of mental phenomena. I contend that there is no evidence for the sort of multiple realizability claimed, but that recognizing this does not undercut the case for autonomous inquiries at higher levels of organization" (p. xi). Craver (2007) also writes: "The different fields that contribute to the mosaic unity of neuroscience are autonomous in that they have different central problems, use different techniques, have different theoretical vocabularies, and make different background assumptions; they are unified because each provides constraints on a mechanistic explanation." (p. 231). And Boone and Piccinini (2016) also discuss traditional "strong autonomy" (related especially with Jerry Fodor), the kind of autonomy they oppose. Given this, I use in this paper the technical term 'explanatory weak autonomy' to clarify *what kind of autonomy* I am discussing. It is not the traditional strong sense, but there is still a sense in which the neo-mechanistic integration includes some weak explanatory autonomy.

for a general explanatory neuroscientific reductive hypothesis on the human neuro-cognitive relationship, based on what neuroscientists are currently doing in the field of molecular and cellular neuroscience.

In this paper, my aim is to investigate whether the neo-mechanistic framework is able to provide a consistent defence of its 'pluralistic integration' in cognitive science in spite of the challenge presented by the neuroscientific neo-reductionist approach. To achieve this goal, firstly, I provide a more detailed characterization of the 21st century mechanistic framework applied to human cognition in cognitive science and discuss its application to the process of memory consolidation (section 1). After this, I characterize the strong neuro-cognitive reductionist approach and its application to the same process of memory consolidation (section 2). The contrast of both approaches shows the challenges the reductionist position brings to the mechanistic one. After the discussion of the divergent points of the two approaches, I analyse the replies offered by the proponents of the neo-mechanistic framework (section 3) and the counter-replies offered by the neuro-cognitive reductionist position (section 4). Based on these analyses, I discuss to what extent the neo-mechanistic framework is successful in its defense of weak explanatory autonomy of cognitive science (section 5). I argue that the framework ultimately incorporates a strong explanatory reductionist position. As a result, any kind of explanatory autonomy for cognitive science is also eliminated.

1 The neo-mechanistic framework for cognitive science

Despite of the many theoretical and terminological differences between the accounts presented by leading neo-mechanists in the field of cognitive science, there are some important common basic points.

The particular application of the general neo-mechanistic account to cognitive science generates a particular theory about scientific activity in the field, which can be called the *mechanistic theory of scientific explanations in cognitive science*. At the same time, this application also generates a theory about the nature of human cognition[3] and of the human neuro-cognitive relationship, which are the most important objects of investigation and explanation in cognitive science. This second theory can be called the *mechanistic theory of human cognition*. These two theories are, evidently, strictly related: the second is concerned with the *explanandum*, and the first with the *explanans*. Together, they provide the neo-mechanistic general framework for investigating human cognition in cognitive science.

[3] Neo-mechanists do not claim that there is something peculiar about human cognition. They assume there is nothing distinctive about it, but this is controversial. Leading cognitive scientists would argue that human cognition has many particular features not found in other cognitive natural organisms, and that aspects of the cognition in these organisms cannot be simply extrapolated to human cognition (cf. BRUNER, 1990; VON ECKARDT, 1993). Thus, I am not using this term because I think neo-mechanists use it, but rather because I think it is more precise and shows more clearly the focus and scope of my discussion.

Thus, the mechanistic theory of human cognition presents an account of what needs to be explained, namely, the human neurocognitive biological mechanism. A biological complex 'mechanism' is normally characterized as "a structure performing a function in virtue of its component parts, component operations, and their organization" (BECHTEL, 2008, p. 13).[4] The core idea is that biological complex mechanisms are systems behaving in a given environment. The general behavior of the whole mechanism is a result of the specific organization of the components and their interactions. According to the mechanistic theory of human cognition, human neuro-cognitive processes are, roughly, information processes performed by neural mechanisms, which represent physical information (cf. BECHTEL, 2008, 2009b; CRAVER, 2007; PICCININI, 2012; THAGARD, 2006; ZEDNIK, 2018). These processes and the mechanisms that perform them can be decomposed in subparts, and these subparts further decomposed. As a result, there can be multiple levels of mechanistic composition in a human neuro-cognitive mechanism. Furthermore, relevant autonomous processes of causation happen in all these different levels (BECHTEL, 2017; CRAVER; BECHTEL, 2007). According to the mechanistic theory of scientific explanations in cognitive science, all these levels and causal processes, in spite of being autonomous, can be related in a pluralistic mechanistic explanation, where the relevant scientific theories are integrated.

One of the clearest examples of an application of the neo-mechanistic framework to a cognitive process is in the domain of memory (cf. BECHTEL, 2008, 2009a; CRAVER, 2007). One important phenomenon related to memory is memory consolidation. Roughly put, this is the phenomenon of transforming short-term memories into long-term memories, what permits the organism to remember important events for a longer period of time and modify its behavior accordingly. To explain this phenomenon, all the relevant regions in the brain responsible for the functions that compose the neuro-cognitive mechanism of memory consolidation, including all relevant mechanistic levels of decomposition, must be identified through the process of localization, i.e. all the particular component parts and component operations of the whole mechanism must be determined.[5] Finally, the causal processes and causal interactions within the functions of the mechanism need also to be understood, i.e. the general organization of the mechanism, and all the different mechanistic levels relevant for the explanation of the phenomenon must be related.

[4] See also the formulations in Craver (2007) and Glennan and Illari (2018).

[5] Firstly, there is the neural systems level which includes for instance "the hippocampus" and its "neuro-architecture" (BECHTEL, 2009a, p. 16, 22). This large neural network (that includes the hippocampus and other particular areas in the brain) is the whole mechanism; and one of the functions that this whole mechanism performs is the phenomenon of memory consolidation, which is the *explanandum* target. The explanation can go further then to a second level, when it decomposes the large neural system into particular sub-neural systems, i.e. a larger neural network that involves larger regions in the brain into smaller neural networks, which are more localized in particular regions. The explanation can go further to another level of decomposition: the inter-cellular level. At this particular level, the components of a particular neural network need to be correctly understood. Finally, the explanation can go to an even lower mechanistic level: the intra-cellular and molecular level. At this level, the description is in terms of the activity of relevant proteins, molecules and ions (cf. BECHTEL, 2009a, p. 18).

2 The strong explanatory neuro-cognitive reductionism

In John Bickle's view, fields such as cognitive neuroscience and cognitive science do not provide the ultimate most complete explanations of human cognition. Instead, he argues that there is an area of current neuroscience, the mainstream of the discipline, which can provide such explanations and is at the same time ruthlessly reductive in spirit: molecular and cellular neuroscience. In his view, at the molecular and cellular lower-level of neural activity very much is already known and it is false that "lower-level neuroscience cannot explain cognition and complex behavior directly." (BICKLE, 2006, p. 411; cf. p. 414). Molecular pathways inside individual neural cells can be linked with cognition and these links "are reductions" (BICKLE, 2008, p. 37).

To accomplish a reduction in Bickle's sense some important steps are needed. The first one is to intervene using molecular genetic methodology into the genome of animals, usually mice. The aim is "to increase or decrease in vivo gene expression and subsequent protein synthesis of intracellular signaling molecules known to be components of pathways that induce and maintain activity-driven synaptic plasticity" (BICKLE, 2012, p. 100). The second step is to measure the effects of the intervention in the behavior of the organism under controlled experimental conditions. The genetically modified animals perform a variety of behavioral tasks so that their specific 'cognitive functions' can be measured. Their behaviors are contrasted with the control animals who are not genetically manipulated. The significant behavioral differences are then understood as the result of the genetic manipulations: the differences in the genetic mechanisms of protein production is the most relevant direct causal factor for explaining the modification in the cognitive function that ultimately produces the behavioral differences (BICKLE, 2012, p. 100).[6]

According to Bickle (2012, p. 101), some scientific experiments already present evidence for establishing the connection between a molecular and cellular mechanism and a particular behavior that indicates a cognitive function. The most clear and detailed example discussed by Bickle is also related to memory, which provides the field of molecular and cellular neuroscience with its "most impressive

[6] The field of molecular and cellular neuroscience has as its central characteristic, thus, the "application of transgenic techniques from molecular genetics into neuroscience. These features allow experimenters to mutate any cloned gene in living, behaving mammals, and thereby manipulate key proteins in intracellular signaling pathways" (BICKLE, 2015, p. 305). These techniques increased the scientific capacity of "manipulation specificity and control" in neuroscience, and generally they "enable a clear picture of not only which neurons have been manipulated but also the specific intracellular signaling pathways affected in those neurons" (BICKLE, 2015, p. 306). In this way, when manipulations of the organisms are successfully made and they produce significant changes in the related behaviors (which can be measured), one can claim that the neural causal-mechanism explains those behaviors, and thereby "the cognitive functions those behaviors are taken to indicate", i.e. "cognitive behaviors" (BICKLE, 2015, p. 305, 306). For the neo-reductionist, research in this area, accordingly, "tests, directly and experimentally, causal-mechanistic hypotheses that purport to explain cognition" (BICKLE, 2015, p. 306). Moreover, in Bickle's view, the development and application of these techniques of engineering genetically mutated mammals, together with the development and applications of the relatively recent technique called 'optogenetics', in order to explain cognition in neuroscience (especially in cellular/molecular neurobiology and behavioral neuroscience) can be considered genuine scientific "revolutions" (2016, p. 1, 2).

achievements" (BICKLE, 2008, p. 36). In one scientific experiment a mouse in which the protein (transcription factor) CREB was 'knocked-out' had intact short-term memory on many rodent memory tasks, while in the long-term memory versions of these tasks there was a great decrease in the memory capacity in comparison with the control. In another experiment CREB was increased in a small population of neurons in a manipulated mouse. This led to an increase of memory consolidation, measured by fear conditioning behavior in the modified mouse. Since CREB has been traditionally considered to be implicated in the induction of late long-term potentiation (L-LTP – a form of neural activity that can last hours, days or weeks which increases neurotransmission efficacy at individual chemical synapses), ultimately is the presence of CREB that is doing, arguably, the bottom-level most central causal work. In fact CREB is part of a particular molecular mechanism that involves cAMP, PKA and CREB, which leads to L-LTP. Bickle claims that blocking any step of this mechanistic process "virtually eradicates memory consolidation, while enhancing steps can lead to faster and stronger consolidation" (2008, p. 38). While these particular experiments are performed mostly in mice, the fundamental hypothesis is that cognition in general can be explained in this way, including human cognition.

The central idea is that this dynamics at the molecular level is actually what *produces* the cognitive processes called long-term memory and memory consolidation. What is being asserted is that there is already sufficient empirical experimental evidence to establish a "causal connection between a proposed cellular or molecular mechanism and a complex, system-level cognitive phenomenon" (BICKLE, 2012, p. 102; cf. SILVA; BICKLE, 2009; SILVA; LANDRETH; BICKLE, 2014).

In order to establish these causal connections, though, Bickle admits that "higher-level scientific investigations" are necessary (BICKLE, 2012, p. 103). This is because precise knowledge about how the whole system behaves is important in order to correlate the "proposed molecular mechanism and the system's behavior we use to indicate the occurrence of a specific cognitive function" (BICKLE, 2012, p. 103). Behavioral experiments at a higher-level also help to establish the "theoretical plausibility of the proposed molecular mechanism for that cognitive phenomenon" (BICKLE, 2012, p. 103-104). Moreover, cognitive neuroscience and its mechanistic goals of decomposition and localization is also important, since it is necessary to identify the most relevant types of neural activity. Consequently, cognitive neuroscience and cognitive science are not dismissed from the relevant and necessary scientific activity in the investigations of human cognition.

Nevertheless, in spite of the necessity of higher level scientific inquire, the "hypothesized molecular mechanism is actually doing the causal work", i.e. "the best causal-mechanistic story for the specific cognitive function then resides at the lowest level of effective experimental interventions." (BICKLE, 2012, p. 104). This means that there is no explanatory pluralism and autonomy here, and this is in sharp contrast with the neo-mechanistic pluralist framework. For Bickle, what counts in the end for delivering the ultimate complete scientific explanation is the

molecular and cellular level. Not all levels are equally significant for the explanation, there is no pluralistic integration, but indeed a reductive one.

3 The mechanists' reply

The neo-mechanistic framework is committed to pluralism: its proponents argue for multilevel causal and explanatory integration in cognitive and neural sciences. Therefore, on the one hand, the framework rejects certain kinds of reduction. However, on the other hand, the framework assumes a kind of 'reductionist' stance. Bechtel claims that "from the point of view of mental activity" his approach is reductionist, and he calls it "*mechanistic reduction*" (2009a, p. 13-14 – highlighted in the original). As he states: "Mechanistic explanation, in seeking to explain the behavior of a mechanism in terms of the operations of its parts, is committed to a form of reduction." (BECHTEL, 2008, p. 129). Mechanistic reduction, in his view, occupies a middle ground between forms of dualism and Bickle's strong reductionism (cf. BECHTEL, 2008, p. 130).

The central problem with strong neuro-cognitive explanatory reductionism (Bickle's reductionism), in Bechtel's view, is due to the fact that "whole systems exhibit behaviors that go beyond the behaviors of their parts" (BECHTEL, 2008, p. 129). Therefore, the strong neuro-cognitive reductionist position is being characterized as defending the thesis that 'individual component parts can explain alone the behavior of a given whole'.

Bechtel, contrary to the extreme reductionist position that he characterizes, wants to be some sort of weak neuro-cognitive explanatory reductionist, since theories related to whole systems are at a higher explanatory level than theories related to their parts, given that wholes are more than their parts. Accordingly, "individual lower-level components do not explain the overall performance of the mechanism. Only the mechanism as a whole is capable of generating the phenomenon" (BECHTEL, 2008, p. 146). Craver's ideas on this point are very similar. In his view, since "mechanisms can do things that individual parts cannot" and "mechanisms explain things that individual parts cannot", thus "higher levels of mechanisms are legitimately included in the explanations of contemporary neuroscience" (CRAVER, 2007, p. 216). As he points out: "Mechanisms require the organization of components in cooperative and inhibitory interactions that allow mechanisms to do things that the parts themselves cannot do" (CRAVER, 2007, p. 216).

In the human neuro-cognitive mechanism of memory consolidation, for example, there are many levels of mechanistic compositional organization, such as the level of connections between systems of neural networks, particular inter-cellular processes, intra-cellular processes, and molecular processes. In each of these compositional levels, there are different causal processes occuring always at the same level, but these causal processes are mediated by the compositional relations, affecting all the different levels of the mechanism (cf. CRAVER; BECHTEL, 2007). Only when all the component parts and their interactions at a lower-level are considered, is it possible to understand the causal activity of the whole

mechanism at the higher-level. The entire neuro-cognitive mechanism is organized in a specific manner, and there are more causal processes occurring at this higher-level (that together are responsible for the behavior of the entire mechanism) than the causal processes at a lower-level related to individual (or sets of) parts. Thus, there is some independent causal higher-level – consequently causal plurality and causal and explanatory weak autonomy. There is here, therefore, causal and explanatory pluralism and weak autonomy, not strong neuro-cognitive explanatory reduction.[7]

Bechtel and Craver intend to claim that identifying components, operations and organization is important to explain the phenomena produced by the whole mechanism. But, at the same time, "the behavior of the whole system must be studied at its own level with appropriate tools for that level", since this level has "a kind of independence" and the phenomena are "different from those studied at the level of the component parts" (BECHTEL, 2008, p. 129). Therefore, the first line of argument is the emphasis on the importance of the internal organization of a given whole, which, in their view, already undermines strong neuro-cognitive explanatory reduction.

Another issue concerning the internal organization of a given whole is related to predictability, and it arises when we consider wholes which are very complex systems. There are systems which are not so complex, as for example, a bike, or a car, or even an airplane. If one wants to understand the functions and behaviors of such systems, all one needs to do is to look to their component parts and subparts to understand how they causally interact with each other. This knowledge about all the components of the system can largely explain what one needs to know about the system, and one can predict much of its behavior in this way, as long as the appropriate knowledge of external conditions is also provided.[8]

In highly dynamical complex systems, there is no linearity in the causal interactions and the components of the system relate not in a static, but in a dynamical way, which makes the system to change constantly its own state and the way it is related to the environment. In such systems accuracy in predictions is not so high. The complex interactions produce new phenomena that are radically different from those related to the components of the system and which cannot be fully predicted or explained just by the knowledge of the operations associated with those components taken in isolation. This occurs because the variables and their interactions are too many and they cannot be measured or understood fully. There will always be a degree of imprecision, a degree of uncertainty in the final value. And since the initial conditions of a system cannot be measured with complete accuracy, i.e. the initial measurement will always have a degree of uncertainty, the results derived from that measurement will also be uncertain.

[7] Craver appears to reject any form of neuro-cognitive explanatory reductionism (2007, p. 228ff.).
[8] To illustrate this, we can think about an airplane under turbulence, or a car being driven in a wet road on a raining day. In such cases, explanations concerning how these simpler systems will behave are often accurate, as well as the related predictions. Evidently, I am not claiming here that there are complete explanatory and one-hundred-percent accurate predictive models for simple systems in cases where the external relevant variables are too many.

As regards neural activity and cognition, if the brain (or parts of the brain) could be considered as such complex dynamical systems (BECHTEL, 2008, 2010), it would be impossible to predict cognitive phenomena arising from them just by looking at physicochemical neural operations and interactions of the components of this brain, or brain region. This is because it would be impossible to know all the initial conditions and to measure all the many variables in the brain responsible for the generation of a certain mental phenomena with total accuracy; it would be too complex. Consequently, any completely accurate prediction of complex human behavior related to mental phenomena would turn out to be impossible. As Craver points out:

> Some mechanisms have so many parts and such reticulate organization that our limited cognitive and computational powers prevent us from making [...] predictions. Some mechanisms are so sensitive to undetectable variations in input or background conditions that their behavior is unpredictable in practice. Behaviors of mechanisms are sometimes emergent in this epistemic sense. (CRAVER, 2007, p. 216-217)

This is another reason why strong neuro-cognitive explanatory reduction fails: certain mechanisms/systems/wholes with highly complex and dynamical internal organization behave in a way that cannot be predicted with high accuracy.

Moreover, as Bechtel emphasizes: "Both the level of the parts and the level of the mechanism engaging its environment play roles in mechanistic analyses. A mechanistic explanation therefore inherently spans at least these two levels." (2008, p. 148). This means that a particular whole mechanism behaves in a certain way "only under appropriate conditions." (BECHTEL, 2008, p. 146). Therefore, the context in which the mechanism behaves is another important aspect for understanding its behavior, not just the internal components, operations and their organization (cf. BECHTEL; ABRAHAMSEN, 2005, p. 426). It is thus crucial for a correct understanding of the behavior of a mechanism to identify the external factors that can affect it:

> No matter how much they investigate the parts, their operations, and their organization, investigators will not identify the variables in the environment that are impinging on the mechanism. Discovering these variables and their effects requires inquiry directed at the environmental variables using appropriate investigatory techniques (BECHTEL, 2008, p. 152).

This factors can be in turn understood as components in a larger mechanism, in which the target mechanism is embedded. It is this mechanistic environment that provides conditions for the rise of the particular behavior. Often, thus, an explanation must clarify what the appropriate environmental conditions for the appearance of a given phenomenon are.

Finally, it is also argued by neo-mechanists in cognitive science that so far there are no clear cases of strong neuro-cognitive explanatory reduction in the field. In their view, even Bickle's most compelling example related to the memory

system does not consist of a top-down search for lower-level mechanisms of memory – Bickle's picture, according to multilevel mechanists, is simply inaccurate and misleading. The example of memory can rather be understood in accordance with the pluralistic view of cognitive neuroscience and cognitive science integration (CRAVER, 2007, p. 237ff). In this view, the LTP neuro-physiological process is considered to be part of the explanation of memory consolidation, not as identical to memory consolidation, nor as a kind of memory consolidation. Accordingly, what Bickle is trying to do is to explain a whole based on its tiny parts. In other words, the particular molecular mechanism that involves cAMP, PKA and CREB, which leads to LTP is, thus, a tiny component of a larger memory mechanism for memory consolidation and cannot alone be considered the ultimate causal explanation for the phenomenon, as Bickle argues. Consequently, there are many important higher-level processes being left out in Bickle's reductive explanation.

Ultimately, therefore, the neo-mechanistic framework for cognitive science stands for pluralistic neuro-cognitive levels of causation and explanation, not for strong explanatory neuro-cognitive reduction.

4 The reductionists' rejoinder

Influential advocates of the neo-mechanistic framework for cognitive science such as Bechtel and Craver argue, thus, for quite basic and straightforward ideas: 1) that the behavior of a whole biological complex mechanism is more than the behavior of its parts taken in isolation; 2) that the organization involving all the components of a mechanism needs to be considered, if one wants to understand the behavior of the entire system; 3) that neuro-cognitive complex mechanisms can be unpredictable to some degree; and 4) that environmental external relations/factors influence the behavior of the entire system. All these points are presented as a case against strong neuro-cognitive explanatory reduction.

The strategy of the advocates of the mechanistic framework is to argue that since the system is organized in a specific manner, and this organization is a high-level aspect of the whole mechanism, because it refers to causal processes that are not just related to individual or sets of parts, but rather to the whole system, then there is some causally independent higher level – consequently, causal and explanatory plurality and weak autonomy.

Indeed, frequently in nature the operation of just one component, or a set of components smaller than the whole mechanism/system, cannot account for the behavior of the whole. For example, the operations of the motor of a car cannot account alone for the whole motion of the car. A whole often has features that none of its constitutive component parts have. This is trivially true, though. Similarly, a collection of *all the component parts* of a whole without considering its organization is also not enough to explain the whole behavior of the mechanism. If one takes just all the component parts of a car without considering how they are related within the car, it is impossible to explain how the car works. In this case, one would be talking about an 'aggregate', not about a whole, since

a mechanistic whole needs to have some organization to be so considered. Therefore, organization including all the components needs to enter the picture.

But Bickle does not argue that a single component's function, or a set of components' functions, always explain the function of a given whole mechanism. This would be a naive view, and Bickle does not take this position. A reductionist position of this kind has never been seriously defended in the specialized literature, because it is clear that the position is trivially false (cf. STEPHAN, 1992, p. 32). The issue here is that what Bechtel and others see as higher-level, for Bickle, can be simply described in terms of lower levels.[9] Bickle knows that organization is important, but in his view all the information about the organization can be also described in terms of lower levels, i.e. at the level of molecules and cells in the case of brain and cognition, where ultimately lies the most important causal mechanistic explanation for him (cf. BICKLE, 2012, p. 104). Theurer and Bickle state that quite directly and explicitly: "[...] as we descend downward through nested mechanisms, opening black boxes at progressively lower-levels along the way, we are uncovering mechanisms that are progressively more explanatory." (2013, p. 106). For Bickle, the organization of all the relevant component parts and their causal interactions can be described according to lower-level terms. The whole mechanism, in his view, is constituted by molecular interactions between a small number of cells, as in the case of the mechanism for LTP. Memory consolidation is a process performed and thus explained by this small molecular mechanism. Therefore, Bickle does not want to take out organization in any sense. This means that the sum of all the component's operations plus their organization in order to provide an explanation for the behavior of the whole mechanism can also accomplish a reduction, as long as all this is described in lower-level science.

Fazekas and Kertész (2011) point out that the possibility of describing all the information at lower levels follows from the very assumptions of the neo-mechanistic framework, because the whole mechanism is just the same as its component parts organized in a particular way. They claim that the mechanistic framework is not able to reduce higher levels to lower levels while maintaining the explanatory weak autonomy of the higher levels simultaneously. In their view, the mechanistic relationship of constitution/composition is not the *asymmetrical* 'part-whole' relation, but the *symmetrical* relation of 'parts + organization = whole', which is identity (FAZEKAS; KERTÉSZ, 2011, p. 373). For them, the ultimate goal of the mechanistic approach "is to explain how a system performs certain tasks by understanding how its parts organized in the right way perform the *very same* task." (2011, p. 372 – highlights in the original). Bechtel and Craver characterize 'lower-level' in terms of individual parts and, then, construct an argument in order to claim that organization is missing at the lower level. This argument against explanatory reduction based on the putative higher-level organization is, however, misleading, because:

> […] it attacks a straw man: it is true (though trivially) if one restricts the characterization of the lower level to the characterization of

[9] Bickle accepts the account of levels of mechanistic composition/parthood as described by prominent neo-mechanists (cf. THEURER; BICKLE, 2013, p. 105).

> individual entities and their behavior. However, we see no reason why the targets of this argument, reductive accounts, should restrict themselves to this 'light' way of characterizing lower-level (FAZEKAS; KERTÉSZ, 2011, p. 377).

There is, therefore, no causal higher levels here, since the same organization can be fully described in lower-level terms. Given that 'parts plus organization' is identical to the mechanism as a whole, the relevant causal explanation for the given phenomenon under investigation is just the same at the lower level, presented with a different vocabulary.[10]

In a similar argumentative line, Soom states that these different mechanistic levels are just "different levels of description" (2012, p. 655), but one can be fully reduced to the other since there is no causal process over and above the organized components of the given mechanism. At each level of description different vocabularies are used by different scientific fields approaching the phenomena at the respective level. But this does not mean that there are novel causal processes at each of these levels. As Rosenberg (2015, § 1) also points out: the neo-mechanists claim that there are higher-levels of explanation because it is possible to identify higher-level autonomous causal processes; at the same time, however, they demand a complete explanation in terms of the organized constitutive/compositional lower-level parts of the mechanism. This is clearly an inconsistency.

Furthermore, Theurer argues that compositional mechanistic relations are transitive, i.e. the information required for the explanation at one level is passed entirely to the other level, since the composition relation of the parts plus organization accounts exhaustively for the behavior of a given mechanism (2013, p. 306). In other words, if a mechanism, M_1, is reduced to a set of sub-mechanisms, M_2, and this set of sub-mechanisms, M_2, is in turn reduced to a set of sub-sub-mechanisms, M_3, then M_1 is directly reduced to M_3. And this occurs in mechanistic explanations because of the ontological and explanatory commitments of the framework: in order to explain the operation of a given mechanism, M_1, one needs to take into consideration all its parts plus organization, i.e. M_2, because M_1 is exhaustively constituted by M_2, as M_2 is exhaustively constituted by M_3. Given this transitive character of mechanistic explanations, lower-levels of molecular and cellular neural mechanisms necessarily explain directly higher-levels of neural

[10] The account of levels presented by neo-mechanists is very different from the account of levels presented by logical empiricists (cf. OPPENHEIM; PUTNAM, 1958). For the latter, theoretical levels of scientific domains are integrated by a deduction of laws. Thus, it is a matter of epistemic deduction: logical, or mathematical. For neo-mechanists, however, levels are integrated by ontic composition of the relevant biological mechanisms. But since these compositional levels can be described in natural language using particular words and concepts, it is correct to say that they can be described with different vocabularies, as neo-mechanists themselves sometimes do. Thus, in biological mechanisms one can consider the dimension of ontic levels of composition as well as the dimension of epistemic levels of description and, consequently, different levels of vocabularies. Moreover, since there is a relationship of complete composition between the whole mechanism and its parts plus organization, the vocabulary related to lower-level parts taken together (not in any isolation whatsoever) will always be referring to the whole too, thereby accounting completely for it in more detail.

systems and cognitive functions, because higher-level explanations become unnecessary after the explanatory work has been done at the lowest-level, which is more detailed and accurate. This is strong explanatory neuro-cognitive reduction. And this is why Theurer and Bickle claim that: "'New' mechanism, by the nature of its key resources, may be far more reductionistic than some of its proponents notice (or admit)." (2013, p. 109 – emphasis in the original).

Even in stochastic cases concerning complex systems, where one cannot have information about all the components due to practical impossibilities leading to unpredictability, there is strong explanatory reduction. There is indeed no conflict between unpredictability in complex systems and Bickle's reductionism. For it is irrelevant if at a certain point in time a phenomenon or a set of phenomena occur that cannot be completely explained or predicted with great accuracy on the basis of the initial state of the system due to the inaccuracy of the measures at the start. What really matters for Bickle's strong neuroscientific explanatory reduction is whether it is possible to describe these phenomena in causal lower-level mechanistic physical language – and then to identify as many mechanistic variables as possible, put all the information together and ultimately produce a scientific explanation of the whole complex system's behavior at hand. In this reductive picture the relevant processes required for the final explanation can be acquired from the lower-level of the components' operations and their interactions; then this information can be put together in order to explain all the causal interactions at the higher-level, since they would be equivalent.

Another line of answer the neo-mechanists provide is to appeal to the external (e.g. environmental, social, cultural) factors that arguably play a causal role at the higher-level of the whole mechanism in determining its behavior. As the argument goes, contextual information is not captured by just investigating the lower-level and its individual parts. In order to capture this information, the higher-level is necessary, and thus there is explanatory weak autonomy (not strong reductive integration) at the higher level. However, there are also counter-arguments for this view. Since the interactions at the higher level are also among mechanisms, i.e. the target mechanism as a whole and its environment, which is composed of other mechanisms, there is nothing that prevents all these whole mechanisms that interact to be further decomposed in sub-components and sub-functions. All of them can be thus described in lower-level vocabulary, as well as their relations to each other. In other words, the reduction to lower levels can be equally applied to mechanisms that interact with a target mechanism as a whole at a given higher level. Fazekas and Kertész (2011, p. 378) correctly argue that everything that is in the higher-level context of a whole target mechanism can be further decomposed and explained in lower levels. Since they are all mechanisms, the reduction should be applied to all of them, without exception.

The argument related to the role culture plays in the development and determination of human cognition is important (cf. BRUNER, 1990; LEITE, 2018, chap. 5), but within the framework of neo-mechanism it is untenable (cf. MILKOWSKI et al., 2018). It is an attempt to formulate, based on a kind of physicalism, a non-reductive explanatory account in similar lines to previous

problematic forms of emergentism (cf. KIM, 2006). All mechanisms are within a larger mechanism (cf. BECHTEL 2008, 2009c; CRAVER, 2007). For instance, the visual system is within the neural system, which is within the biological system, which is within an ecological system, and so on. In this way, all causal factors external to the target mechanism are causes from parts that belong to a larger mechanism, which in turn can also be decomposed and understood at the lower level. Moreover, since all mechanisms are completely composed by all the lower-level parts plus organization, and the internal causes are mediated by the compositional relation, the effects of such external causes will be felt at this lower level, and can be explained at this level if all the important parts are considered. For example, if particular areas of the brain affect the area V1 in the occipital lobe, these effects can be understood at the level of cells and molecules, as neuroscientific research has been constantly showing. The same kind of decomposition can be applied to these other particular areas, external to V1. In fact, the 'higher level' in mechanisms is merely an abstraction, if we are not considering just some part of the mechanism. In this case, there is no 'higher level'. Referring to a 'higher level', after considering all the lower level parts plus the organization, is simply misleading.

Finally, the neo-mechanists' arguments appealing to the fact that allegedly there are no real cases of neuro-cognitive reduction in the cognitive and neural sciences are not going to convince Bickle and other neo-reductionists. Bickle can simply maintain that his theory of strong neuro-cognitive explanatory reduction concerning the capacity of memory consolidation provides a plausible case where neuro-cognitive reduction occurs in the cognitive and neural sciences. He can also argue that this kind of reduction is not peripheral and that it is, on the contrary, ubiquitous in neuro-cognitive science. Therefore, those arguments won't succeed in order to undermine Bickle's position and save neo-mechanistic neuro-cognitive pluralist integration.[11]

5 The problem for neo-mechanists in cognitive science remains

This contemporary controversy between the neo-mechanists and the neo-reductionists has implications for establishing theoretical foundations and theoretical integration in the contemporary field of cognitive science (including here cognitive neuroscience). Therefore, it must be considered carefully by those interested in this scientific area.

The first important issue to consider in this regard is concerned with the *explanans*. The question here is whether the authors on these two sides have a common view on what successful scientific explanations in cognitive science are. Bechtel and Craver are leading neo-mechanists working in the field of

[11] It is important to note that, in this paper, I am pointing out theoretical problems. Empirical research is extremely useful to discuss them, but they will not solve these problems alone. Besides, there is already in the literature a substantial discussion of empirical research. Now it is important to discuss more deeply the theoretical problems. I do not think, therefore, that more empirical research will be useful to settle this issue, and thus it is of little help to add more.

neuroscience and cognitive science. They have been articulating in the last decades a view of scientific explanations for these and other fields of contemporary science (BECHTEL; ABRAHAMSEN, 2005; BECHTEL, 2008; MACHAMER; DARDEN; CRAVER, 2000; CRAVER, 2007). Bickle and other neo-reductionists are equally making claims about how explanations in neuroscience should be constructed (BICKLE, 2006; BICKLE; THEURER, 2013). While their accounts may differ on important aspects, all these authors agree with some basic notions. For instance, none of them subscribe to a view of scientific explanation based on logical derivation of laws and theories in a deductive argument – a model defended most prominently by some logical empiricists around the middle of the last century.

For these contemporary authors, properly understanding and explaining some natural biological phenomenon is a matter of providing a description of the biological mechanism responsible for generating this phenomenon. To explain here, thus, is to provide descriptions of working components and subcomponents of macro-mechanisms, their regular and irregular interactions and how the macro-mechanism is affected by external varying conditions. The macro-mechanism is considered to produce the phenomenon under investigation, i.e. what ultimately needs to be explained. The most accurate and detailed the account of the macro-mechanism and the external conditions that affect its functioning are, the better, because this will be more explanatory. Thus, there is a minimum common ground for comparing the explanatory frameworks in cognitive science theoretically, even if they present different views on some particular topics.

The second issue is related to the *explanandum* and to empirical evidence about its explanation. The question here is whether the authors in the debate are using examples of explanations in cognitive science concerned with the same phenomena, or very similar phenomena, that can be easily compared. This is indeed not the case. Bechtel (2009a) uses examples of episodic memory consolidation in humans, Craver (2002, 2007) uses the example of spatial memory consolidation in mice, and Bickle (2012) uses the example of memory consolidation for fear conditioning in mice. It would be certainly helpful if we could make these discussions and comparisons more systematic by using the same examples. In this way, one would be able to understand more precisely to what extent the different theoretical frameworks can be successfully applied in order to explain particular empirical phenomena, and which one presents a better explanation (cf. THEURER, 2013). Otherwise, while Bechtel and Craver will claim that the macro-mechanism amounts to *A*, *B* and *C*, Bickle will claim that the macro-mechanism amounts actually to *A*. In this case, the attempts of comparison will start by begging the question and reach nowhere. Nevertheless, apart from the difficulties in comparing empirical results to make a case for a particular framework, there is a more theoretical issue that remains a problem for the leading neo-mechanists.

Based on the discussions analyzed in the previous sections, one can note that the crucial point of disagreement is related to the relationship between levels in the mechanistic framework and which level is more explanatory. Bechtel and Craver (2007) claim that their account is 'causal at the intra-level' and 'constitutive

at the inter-level'. This is the so called 'mechanistic constitution' relation (CRAVER, 2007). In a recent publication, Povich and Craver state that "constitutive mechanistic explanations" are those in which "one can explain the behavior of the whole in terms of the organized behaviors of its parts, and one can explain the behaviors of the parts in terms of the organized behaviors of their parts." (2018, p. 186). They repeat that "the whole has capacities the parts alone do not possess" (2018, p. 190) and they say, furthermore, that: "If two things are not related as part to a whole, they are not at different levels" (2018, p. 188).

However, if one defines lower-level just in terms of what a part does and state, at the same time, that there can be no complete explanation of the behavior of a mechanistic whole based just on what parts do, then it is already *in principle wrong* to provide an explanation based on the lower level. The point articulated by the neo-reductionists is just that the lower level does not need to be defined in these terms. It is rather a matter of looking deeper into the sub-parts of the parts and provide an explanation based in that vocabulary, at that level of description, with all the required details. All the organization of the macro-mechanism, together with the behaviors of all the relevant parts, can have such description, at least in some cases.

Moreover, the mechanistic constitution idea appears indeed to lead to an identity relationship, as Fazekas and Kertész (2011) argue, since the mechanists are committed to the idea of explaining *completely and exhaustively* the function of a mechanism by its constitutive parts and organization in a given context. Povich and Craver state that: "The behavior of the whole contains the behaviors of the parts, and the behaviors of the parts collectively and exhaustively constitute the behavior of the whole." (2018, p. 193). If this is so, then there is no further space for anything at some higher level that cannot be described using a vocabulary at a lower level. Anything that is necessary for the explanation can be fully described:

> If one knows all of the relevant entities, activities, and organizational features, and knows all the relevant features of the mechanism's context of operation, and can in principle put it all together, then one must know how the mechanism will behave. […] It is an epistemic warning sign if features of the mechanism's behavior cannot be accounted for in terms of our understanding of its parts, activities, organization, and context. Mechanists thus operate with a background assumption that the phenomenon is exhaustively explained (in an ontic sense) by the organized activities of parts in context (POVICH; CRAVER, 2018, p. 193).

However, Povich and Craver claim that given the possibility of multiple realizability, token and type identity are not good terms to characterize their mechanistic relationship between levels (2018, p. 193). Another problem that they point out is that we do not have criteria to establish if a psychological function is indeed identical to another; thus, it is very hard to know if a psychological function is indeed identical to a function performed by a given mechanism. If it is so, then, this is as much a problem for 'identity relations' as it is for 'complete and exhaustive mechanistic constitution'. For if a psychological function, F_1, can be performed by

different macro-mechanisms, M_1, M_2 and M_3, then just providing a description of all the relevant activities related to M_1 will not be enough to provide an exhaustive description of the production of F_1. In case the number of macro-mechanisms responsible for F_1 is countable and their activities explainable, an identification between the psychological function and the number of mechanisms producing it is still possible. In case the number is uncountable, there can be no exhaustive constitution and thus explanation for that psychological function. The same happens if the psychological function cannot be identified as being the same function in different times considering at least the properties that are most relevant for that characterization. In this case, it is not possible to claim that a given macro-mechanism completely and exhaustively constitutes that psychological function, because there could be no certainty about that, or high probability that it is the case.

However, in case exhaustive mechanistic constitution is plausible, there is no good reason to suppose that the internal macro-organization of a whole human neuro-cognitive mechanism cannot be explained by lower-level components and their interactions, at least in some cases, such as memory consolidation as described by some neo-mechanists in neurophysiological terms. In other words, there is no reason to suppose that an account of the organization of the parts of a given macro-mechanism in some cases cannot be provided using the vocabulary of the lower level. The causal novelty or independence that is supposed to appear at the higher level is never clearly pointed out. But if it is present in the model, then one cannot claim that the behavior of the whole macro-mechanism will be entirely and exhaustively explained by all its component parts plus their organization in a particular context.

Someone defending the neo-mechanistic framework might still try to argue that the organization of a macro-mechanism can only be characterized with higher-level vocabularies, since the organizational feature of a mechanism is more than the spatial disposition of its parts: it includes complex causal interactions at different times between many operations and activities.

However, what I am showing in this paper is precisely that this argument is untenable. All the features related to the internal organization of biological mechanisms are compositional in nature. Consequently, as soon as they are completely decomposed by empirical research and all these decomposed parts and their operations are properly related in lower-level vocabulary, a complete account of the phenomenon under investigation will be provided at this lower level. There is no way out of this. Any attempt to show that there is a significant novelty at a supposed higher level will necessarily damage the mechanistic integration by exhaustive composition.

The source of the problem here is the emphasis from leading neo-mechanists in considering just parts in isolation when talking about lower levels. This creates the illusion that lower levels are always parts in isolation. But the lower levels of mechanistic composition are, on the contrary, all the parts and operations fully decomposed and taken together, i.e. organized together. Consequently, the organization is not just at a 'higher level'. It can be understood

and described at the lower level as well, because this organization (even if it is a very complex one) is simply the interactions between the parts (composed or fully decomposed) that exhaustively constitute the whole mechanism. Thus, even the most complex organization can be fully described in a lower-level vocabulary, as the research in the field of molecular and cellular neuroscience shows.

There is an internal contradiction within the neo-mechanistic framework for cognitive science: a full explanation of a given neuro-cognitive phenomenon requires all the mechanistic components plus the organization to be described, what can be completely and exhaustively done at a lower level, but if there is some novel and autonomous causal element at a higher level (given some sort of ontological emergence, or multiple realizability), no full explanation is possible at the lower level. This problem of the *inconsistence and instability in neo-mechanistic integrative explanations* is the central problem that remains as an obstacle for the pluralist integration defended by many supporters of the neo-mechanistic framework in cognitive science.

Concluding remarks

The neo-mechanistic framework has been very influential in cognitive science, as well as very significant for many important issues intensively debated in the field, such as those related to scientific integration.

However, the most central arguments the advocates of the framework use against strong neuro-cognitive explanatory reduction present significant shortcomings. Neo-mechanists argue that the strong neuro-cognitive reductionist is committed to the view that 'a part of a whole must explain the behavior of the entire whole'. That is not correct, however. Bickle is aware that individual parts do not explain the behavior of the whole in the life sciences. Instead, what he claims is that all the aspects of such explanation can be described at a lower level. This includes the organization of the whole biological neuro-cognitive mechanism, biological neuro-cognitive mechanisms with high degrees of complexity, and the external factors affecting the whole biological neuro-cognitive mechanism.

I am not presenting a new defense of neo-reductionism here. Bickle's framework is not being advocated here as the most plausible account for investigating human cognition in cognitive science. The strong neuro-cognitive reductionist framework presents indeed many problems (cf. e.g. LOOREN DE JONG; SCHOUTEN, 2005; LOOREN DE JONG, 2006). But it was not my purpose to analyse the plausibility of this neo-reductionist framework as a whole in this paper.

I am simply taking the view that neo-mechanists' defense of 'weak reductionism'/ 'weak pluralist integration' is inconsistent. The neo-mechanistic framework for the cognitive and neural sciences gives full space for Bickle to advance his approach. Underlying the neo-mechanistic framework is a strong neuro-cognitive reductionist program. This can be seen through the examination of the ultimate consequences of the theoretical commitments presented by

influential neo-mechanists. And Bickle is, evidently, taking full advantage of this (cf. THEURER; BICKLE, 2013).

The analysis shows, therefore, that, although the neo-mechanistic framework for cognitive science is allegedly committed to neuro-cognitive causal and explanatory pluralism, it is not able yet to provide a consistent defense of it, collapsing into a strong neuro-cognitive explanatory reductionist framework, or being an inconsistent framework. As a result, the neo-mechanistic framework's defense of neuro-cognitive explanatory pluralist integration is untenable. No robust, meaningful, or even weak explanatory autonomy of cognitive science can be achieved in this way.

References

BECHTEL, William; ABRAHAMSEN, Adele. Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, v. 36, p. 421-441, 2005.

BECHTEL, William; RICHARDSON, Robert. *Discovering complexity*: decomposition and localization as strategies in scientific research. Cambridge, MA: MIT Press, 2010. (Originally published in 1993).

BECHTEL, William; WRIGHT, Cory. What is psychological explanation? In: SYMONS, John; CALVO, Paco (eds.). *The Routledge Companion to Philosophy of Psychology*. New York: Routledge/Taylor & Francis Group, 2009. p. 113-130.

BECHTEL, William. Levels of description and explanation in cognitive science. *Minds and Machines*, v. 4, p. 1-25, 1994.

_____. Reducing psychology while maintaining its autonomy via mechanistic explanations. In: SCHOUTEN, Maurice; LOOREN DE JONG, Huib (eds.). *The matter of the mind*: philosophical essays on psychology, neuroscience, and reduction. Malden, MA: Blackwell Publishing, 2007. p. 172-198.

_____. *Mental mechanisms:* philosophical perspectives on cognitive neurosciences. New York: Routledge, 2008.

_____. Molecules, systems, and behavior: Another view of memory consolidation. In: Bickle, John (ed.). *The Oxford Handbook of Philosophy and Neuroscience*. New York: Oxford University Press, 2009a. p. 13-40.

_____. Constructing a philosophy of science of cognitive science. *Topics in Cognitive Science*, v. 1, n. 3, p. 548-569, 2009b.

_____. Looking down, around and up: mechanistic explanations in psychology. *Philosophical Psychology*, v. 22, n. 5, 543-564, 2009c.

_____. How can philosophy be a true cognitive science discipline? *Topics in Cognitive Science*, v. 2, p. 357-366, 2010.

_____. Explicating top-down causation using networks and dynamics. *Philosophy of Science*, v. 84, p. 253-274, 2017.

BICKLE, John. *Philosophy and neuroscience*: a ruthlessly reductive account. Norwell, MA: Kluwer Academic Press, 2003.

_____. Reducing mind to molecular pathways: explicating the reductionism implicit in current cellular and molecular neuroscience. *Synthese*, v. 151, p. 411-434, 2006.

_____. Real reduction in real neuroscience: metascience, not philosophy of science (and Certainly Not Metaphysics!). In: HOHWY, Jakob; KALLESTRUP, Jesper (eds.). *Being reduced*: new essays on reduction, explanation, and causation. Oxford: Oxford University Press, 2008. p. 34-51.

_____. A brief history of neuroscience's actual influences on mind–brain reductionism. In: GOZZANO, Simone; HILL, Christopher S. (eds.). *New perspectives on type identity*: the mental and the physical. Cambridge: Cambridge University Press, 2012. p. 88-110.

_____. Marr and reductionism. *Topics in Cognitive Science*, v. 7, p. 299-311, 2015.

_____. Revolutions in neuroscience: tool development. *Frontiers in Systems Neuroscience*, 2016. doi: 10.3389/fnsys.2016.00024.

BOONE, Worth; PICCININI, Gualtiero. The cognitive neuroscience revolution. *Synthese*, v. 193, n. 3, p. 1509-1534, 2016. doi:10.1007/s11229-015-0783-4.

BRUNER, Jerome. *Acts of meaning*. Cambridge, MA: Harvard University Press, 1990.

CRAVER, Carl F.; BECHTEL, William. Top-down causation without top-down causes. *Biology and Philosophy*, v. 22, p. 547-563, 2007.

CRAVER, Carl F.; TABERY, James. Mechanisms in Science. In: ZALTA, Edward N. (ed.). *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), 2015. Available at: https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/.

CRAVER, Carl F. Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, v. 69, n. S3, p. S93-S97, 2002.

_____. *Explaining the brain*: mechanisms and the mosaic unity of neuroscience. Oxford: Oxford University Press, 2007.

FAZEKAS, Peter; KERTÉSZ, Gergely. Causation at different levels: tracking the commitments of mechanistic explanations. *Biological Philosophy*, v. 26, p. 365-383, 2011.

GLENNAN, Stuart; ILLARI, Phyllis (eds.). *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. London: Routledge, 2018.

KAPLAN, David. Neural Computation, Multiple Realizability, and the Prospects for Mechanistic Explanation. In: KAPLAN, David (ed.). *Explanation and integration in mind and brain science*. Oxford: Oxford University Press, 2017. p. 164-189.

KIM, Jaegwon. Emergence: core ideas and issues. *Synthese*, v. 151, n. 3, p. 347-354, 2006.

LEITE, Diego A. *The twenty-first century mechanistic theory of human cognition:* a critical appraisal. Doctoral Dissertation. University of Trento (Italy), 2018. Available at: http://eprints-phd.biblio.unitn.it/2828/.

LOOREN DE JONG, Huib; SCHOUTEN, Maurice. Ruthless reductionism: a review essay of John Bickle's philosophy and neuroscience: A ruthlessly reductive account. *Philosophical Psychology*, v. 18, n. 4, p. 473-486, 2005.

LOOREN DE JONG, Huib. Explicating pluralism: where the mind to molecule pathway gets off the track – reply to Bickle. *Synthese*, v. 151, p. 435-443, 2006.

MACHAMER, Peter; DARDEN, Lindley; CRAVER, Carl F. Thinking about mechanisms. *Philosophy of Science*, v. 67, n. 1, p. 1-25, 2000.

MILKOWSKI, Marcin. Integrating cognitive (neuro)science using mechanisms. *Avant: Journal of Philosophical-Interdisciplinary Vanguard*, v. VI, n. 2, p. 45-67, 2016. doi:10.26913/70202016.0112.0003.

MILKOWSKI, Marcin et al. From wide cognition to mechanisms: A silent revolution. *Frontiers in Psychology*, v. 9, 2018. doi:10.3389/fpsyg.2018.02393

OPPENHEIM, Paul; PUTNAM, Hilary. The unity of science as a working hypothesis. In: FEIGL, Herbert; SCRIVEN, Michael; MAXWELL, Grover (eds.). *Concepts, theories, and the mind-body problem*. Minneapolis: Minnesota University Press, 1958. p. 3-36.

PICCININI, Gualtiero; BAHAR, Sonya. Neural computation and the computational theory of cognition. *Cognitive Science*, v. 37, n. 3, p. 453-488, 2013.

PICCININI, Gualtiero; CRAVER, Carl F. Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, v. 183, n. 3, p. 283-311, 2011.

PICCININI, Gualtiero. Computing mechanisms. *Philosophy of Science*, v. 74, n. 4, p. 501–526, 2007.

_____. Computationalism. In: MARGOLIS, Eric; SAMUELS, Richard; STICH, Stephen P. (eds.). *The Oxford Handbook of Philosophy of Cognitive Science*. New York: Oxford University Press, 2012. p. 222-249.

_____. *Physical computation*: a mechanistic account. Oxford: Oxford University Press, 2015.

POVICH, Mark; CRAVER, Carl F. Mechanistic levels, reduction and emergence. In: GLENNAN, Stuart; ILLARI, Phyllis (eds.). *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. London: Routledge, 2018. p. 185-197.

ROSENBERG, Alex. Making mechanism interesting. *Synthese*, 2015. doi: 10.1007/s11229-015-0713-5.

SILVA, Alcino; BICKLE, John. The science of research and the search for molecular mechanisms of cognitive functions. In BICKLE, John (ed.). *The Oxford Handbook of Philosophy and Neuroscience*. New York: Oxford University Press, 2009. p. 91-126.

SILVA, Alcino; LANDRETH, Anthony; BICKLE, John. *Engineering the next revolution in neuroscience*: the new science of experiment planning. New York, NY: Oxford University Press, 2014.

SOOM, Patrice. Mechanisms, determination and the metaphysics of neuroscience. *Studies in History and Philosophy of Science*, v. C43, p. 655-664, 2012.

STEPHAN, Achim. Emergence – a systematic view on its historical facets. In: BECKERMANN, Ansgar; FLOHR, Hans; KIM, Jaegwon (eds.). *Emergence or reduction? Essays on the prospects of nonreductive physicalism*. Berlin: Walter de Gruyter, 1992. p. 25-48.

THAGARD, Paul. *Hot thought: mechanisms and applications of emotional cognition*. Cambridge, MA: MIT Press, 2006.

_____. Why cognitive science needs philosophy and vice versa. *Topics in Cognitive Science*, v. 1, p. 237-254, 2009.

_____. Cognitive Science. In: ZALTA, Edward N. (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), 2018. Available at: https://plato.stanford.edu/archives/win2018/entries/cognitive-science/.

THEURER, Kari; BICKLE, John. What's old is new again: Kemeny-Oppenheim reduction in current molecular neuroscience. *Philosophia Scientiae*, v. 17, n. 2, p. 89-113, 2013.

THEURER, Kari. Compositional explanatory relations and mechanistic reduction. *Minds & Machines*, v. 23, p. 287-307, 2013.

VON ECKARDT, Barbara. *What is cognitive science?* Cambridge, MA: MIT Press, 1993.

WRIGHT, Cory; BECHTEL, William. Mechanisms and psychological explanation. In: THAGARD, Paul (ed.). *Philosophy of psychology and cognitive science*. Amsterdam: North Holland Elsevier, 2007. p. 31-79.

ZEDNIK, Carlos (2018). Mechanisms in cognitive science. In: GLENNAN, Stuart; ILLARI, Phyllis (eds.). *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. London: Routledge, 2018. p. 389-400.

Recebido em: 14-01-2019

Aceito para publicação em: 25-06-19