A Cautionary Contribution to the Philosophy of Explanation in the Cognitive Neurosciences

Abstract

I propose a cautionary assessment of the debate concerning the impact of the dynamical approach on philosophical accounts of scientific explanation in cognitive neuroscience. I criticize the dominant mechanistic philosophy of explanation, pointing out that it doesn't do justice to the field's diversity and stage of development, and that it fosters misguided interpretations of dynamical models' contribution. In order to support these arguments, I analyze a case study in computational neuroethology and show why it should not be understood through a mechanistic lens. In addition, I argue for a greater appreciation of the relation between explanation and other epistemic goals.

Key words: Explanation; dynamical approach; scientific models; mechanisms; philosophy of neuroscience.

1. Introduction

In recent years the problem of scientific explanation has gained a renewed and unprecedented importance in the philosophy of cognitive science and neuroscience. The debate about the kind of explanation that is pursued in these fields has been dominated by the idea of a mechanistic explanation, a position that has its origin in the philosophy of biology of William Wimsatt and has been consolidating in the last decade. Such has been the growth of this position that you can often read of a "new mechanistic philosophy of science" (for example, Bechtel and Richardson 2010: xvii), which articulates the general thesis that in certain scientific fields, and specially in the biological sciences, phenomena are explained through the identification and specification of the mechanisms responsible for them.

While the mechanistic position has been initially elaborated with an eye on molecular and cell biology, it has also been, and is today more than ever, the dominant position in the case of cognitive science and specially of cognitive neuroscience. A number of philosophers, among which William Bechtel, Carl Craver and Peter Machamer can be mentioned, have developed and applied it to a number of successful research cases. Setting aside the differences between the different authors, there is a common core that shapes the mechanistic position, exemplified as it has been by results from research on several cognitive phenomena, such as memory and visual perception, as well as of various neural phenomena, like maintenance of circadian rhythms or action potential generation. The associated defense of this position has responded, on the one hand, to variants of functionalism for a number of areas in cognitive science (for example, Weiskopf 2011) and, on the other, and increasingly so in recent years, to the growing visibility of the so-called dynamical approach to cognitive science (van Gelder and Port 1995; Beer 2000).

The latter front of debate deserves special attention, in part because it has taken a life of its own and also because it has somehow gathered much of the philosophical reflection on this line of work –in a similar way in which the debate over the need of mental representations had during the late past century. In this respect, the number of recent publications specifically dedicated to the tension between the mechanistic model of explanation and dynamical research is noteworthy: Bechtel (2001), van Leeuwen (2005), Chemero and Silberstein (2008), Walmsley (2008), Kaplan and Bechtel (2011), Kaplan and Craver (2011), Stepp, Chemero, and Turvey (2011), Zednik (2011), Abrahamsen and Bechtel (2012), Silberstein and Chemero (2013), Gervais (2015), Ross

(2015), among others. The extensive use, typical of the dynamical approach, of nonlinear dynamical systems theory as a tool for developing models and analyzing experimental data seems to question some important assumptions about the kind of explanation, and therefore the kind of theories, which must be sought in this field of scientific research. According to some philosophers, following considerations hinging on the mathematical apparatus adopted for model building, it appears to be emerging a kind of explanation distant from the one behind more established results, conceived in terms of mechanisms.

I here lay some bases for a to some extent negative forecast on the developing debate on how this line of research will impact on our understanding of explanation. To this end, I will focus on the arguments of those who defend the mechanistic model of scientific explanation as applied to dynamical research as well as some of those opposing it. At the same time, I suggest some guidelines so as to highlight those aspects of the debate that are relevant to the planning and development of research in this complex scientific field. Thus, I intend to make a cautionary contribution primarily directed to ponder the debate's role and relevance. Although I will consider some specific aspects of the conflicting positions, the kind of proposals I will develop below are largely external to them.

My main argument hinges on an appraisal of the level of development and the complexity of practices that characterize cognitive science and, particularly, contemporary cognitive neuroscience. My central general point is that this is a case where the field in question is not mature enough for the associated philosophical debate, as it is laid out in the literature. On the other hand, I show that a more diachronic view of neuroscience research, conceived in terms of problem solving, or answers-to-questions dialectic, reveals that only some of the questions that stimulate scientific research concern directly explanation. Accordingly, I favor a shift in perspective on the problem of scientific explanation in this area, in a way that it is framed within a context of multiple epistemic goals, more or less relevant depending on the particular case under consideration.

It should also be noted that an important motivation for the present contribution is that of strengthening a more general cautious attitude. It thus points to the philosophy of cognitive science and neuroscience as emerging philosophical areas where there lacks a strong tradition that could serve as a standard regarding the kind of approach and the scope of the existing problems. In this way, I'll show that it is necessary to beware of the application of debates pertaining to general philosophy of science to the young field of cognitive neuroscience and related disciplines. Hence the importance of the projection and proposed redirection of the debate I articulate in what follows.

The structure of the paper is as follows. I start presenting the tension that dynamical models generated as they proliferated in the cognitive sciences, a tension raised in terms of their explanatory value. Here I focus mainly on the debate between those who interpret these models as part of mechanistic research and those who oppose this possibility. Then I develop my critical stance towards certain aspects of the presented debate: First, I offer some general considerations about the heterogeneity and the state of development of cognitive neuroscience, and how these issues should impact on the problem of explanation in this field. Then I show in what sense mechanistic positions confronting dynamical research lead to a number of undesirable consequences: Namely, the extent to which its theoretical contribution is misunderstood, how its interpretation is forced into mechanistic terms, and a lack of appreciation of their exploratory nature. To illustrate my claims, I take Zednik (2011) as my disputant and, in particular, his proposal for an expanded mechanicism for certain cases of dynamical research in cognitive neuroscience. Finally, I question the dominant role of the topic of explanation, particularly in its mechanistic variant, in the philosophical reflection on the scientific area of interest.

2. Dynamical systems, mechanisms, and explanation in cognitive science

It would not be an overstatement to say that the philosophy of cognitive and brain sciences is strongly dominated by the problem of scientific explanation. This is a situation that has been consolidating within the philosophical community at least since the turn of the century. In particular, the mechanistic model, a particular kind of explanatory model proposed as a solution to this problem (for example, Glennan 2002; Craver 2006; Bechtel 2008; just to name some relevant accounts), has witnessed a phenomenal development in the field. The philosophical debate about the idea of mechanism conceived as the foundation for explanation has taken a life of its own also when it comes to its connection with the so-called dynamical approach (Chemero and Silberstein 2008; Kaplan and Bechtel 2011; Kaplan and Craver 2011; Zednik 2011; Silberstein and Chemero 2013; Gervais 2015; Ross 2015; among others). The problem here lies in the scope of mechanistic models in cognitive science, and the compatibility between them and dynamical models.

In this context, dynamical models are those models applied to the study of cognitive processes and their neural basis, whose development depends centrally on the use of the mathematical theory of nonlinear dynamical systems (for example, Strogatz 1994). A number of philosophers and scientists articulated, in different ways, the idea that this extensive application results in a characteristic approach to modeling and experimental data analysis (van Gelder and Port 1995; Beer 2000; Chemero and Silberstein 2008; Schöner and Reimann 2008; among others). In this sense, the field of interest has witnessed an increasing adoption of concepts, analytical methods and graphical tools provided by this theory, which allows to conceptualize and model complex systems in their temporal evolution. Generally speaking, the theory is used for the purpose of guiding research, in different ways, in disparate areas of the cognitive and brain sciences, a trend that has led to the idea of a dynamical line of work, a dynamical approach.

The application of the concepts and methods of dynamical systems theory affords addressing cognitive and neural phenomena in terms of trajectories in a space of possible states of a previously defined system –usually through systems of differential equations–, as well as analyzing and visualizing the course of its temporal evolution. The obtained models tend to characterize the system's behavior in terms of higher order variables that describe some of its global states, which in turn grants them a rich descriptive power. The benefits of this qualitative perspective lay partly in the universality of the properties used to characterize the analyzed dynamical systems (for example, kinds of couplings between oscillators, of instabilities, attractors, bifurcations, and so on). It has been stated in this regard that the dynamicist explanatory focus is set on the structure of the space of possible trajectories of the system under study and the internal and external forces that shape them (cf., Beer 2000: 96).

The potential application range of this modeling style is worth highlighting. Dynamical language enables the construction of models that integrate disparate aspects of the phenomenon under study, with an emphasis on their temporal unfolding, on different time scales. Many researchers see in this mathematical language the possibility to address multiple simultaneous influences within a system, often characterized by different change rates, that ultimately result in the system's behavior. As expressed by Smith and Thelen (cf., 2003: 344), the dynamics of a time scale are continuous with, and are nested within, the dynamics of other time scales (for example, scales of neuronal excitability, action, learning, development, evolution) and it is essential to capture the temporal interactions characteristic of behavior within and between different scales.

Crucially, this breadth of scope allows for the inclusive consideration of aspects pertaining to the cognitive system (interpretable in neural, cognitive or behavioral terms) as well as specific aspects of the agent's body and environment, according to their respective differing time scales. This has in fact led to a proliferation of dynamical models in the context of research of the most diverse phenomena and under any further methodological restriction: Including more or less qualitative, more or less quantitative variants (distinguishable, for example, in terms of the accuracy of predictions they allow), and a variety of neural network models. (I shall return later to this issue in a critical vein vis-à-vis the problem of explanation.)

In the case of cognitive neuroscience, we can also talk of dynamical lines of work, that is, those research programs that prioritize the use of dynamical systems theory to address several aspects of dynamic patterns formation in the brain. Part of the appeal here is the ability to compress the very high number of degrees of freedom of (a specific delimitation of) a system under consideration to some variables that manage to capture global and temporarily extended features of neural dynamics and that at the same time can be connected to cognitive performance. Again, the temporal dimension here takes on a central role, to the detriment of the consideration of a number of finer grain structural variables. Some relevant examples of this kind of work are Engel, Fries, and Singer (2001), Freeman (2005), Izhikevich (2007) and Buzsaki (2011).

Now, despite its continued development, the widespread reception of the dynamical approach has been largely negative, so much so in its inaugural years by the mid-nineties as, though to a lesser extent, in more recent years. Recently, the explanatory credentials of this kind of approach have been a particularly common front of attack by philosophers and scientists. In the literature, we can distinguish two different but complementary critical strategies, which have been used in this direction.

One of the hardest attacks from some of these critics (Dietrich and Markman 2001; Bechtel 2001; van Leeuwen 2005) has been to bluntly deny the explanatory character of dynamical research due to reasons inherent to the application of dynamical systems theory. Along this line, the mathematical language of dynamical systems theory, which can be seen as the dynamical approach's operational heart, is nothing more than a good descriptive, not explanatory, tool for addressing the behavior of complex systems. If we consider the nonlinear equations used to elaborate the models, the central idea is that, if their terms are referentially opaque –inasmuch as variables describe global states of the modeled system–, all that dynamical research can offer are sophisticated redescriptions, but never explanations, of phenomena (for an analysis and response regarding this strategy, see Author 2011). The notion of mechanism is, as anticipated, the protagonist of the second wave of assaults against the explanatory potential of dynamical research. I turn now to this strategy.

It should be noted at the outset that this attack is complemented by the one previously presented to the extent that, on the one hand, the dynamical approach is considered insufficient as an autonomous approach in cognitive science and, on the other hand, it is made subsidiary to strategies aimed at uncovering cognitive and neural mechanisms. To the extent that it stands as a characteristic approach that exceeds strictly neuroscientific research, it is necessary to be clear about the relative independence between the problem of explanatory format, on the one hand, and issues related to reductionism, on the other. This point is relevant because, although this approach was first explored in neuroscience –where Walter Freeman's program in neurodynamics is a pioneering example–, it then grew strong also in areas such as perception theory, cognitive robotics, or developmental psychology.

Arguably, both critiques confronting dynamical research to some extent have as background a certain philosophical tradition surrounding explanation in cognitive science, from which specific positions emerge regarding the question of reductionism. In general, according to this tradition, an explanation in cognitive science accounts for a system's behavior in terms of the causal interaction between its component parts and their relative contribution to the information processing necessary for the task under consideration. A very influential development in this direction has been to frame this kind of explanation within the double heuristic of decomposition and localization, mostly due to Bechtel and Richardson's (2010) classical account.

What's relevant to this particular way of framing the problem of explanation is that it provides a common space to shelter opposing positions with regard to the autonomy of psychological theorizing and the related problem of reductionism in the cognitive and brain sciences. As Chemero and Silberstein (2008) argue, the problem for the dynamical approach is whether its explanations must be understood as mechanistic, while only a positive answer to this question would lead us to the question of reductionism. What the problem inquires, then, is just whether a system's behavior can be accounted for in terms of the functions carried out by its parts and the organization of their interaction.

According to the mechanistic tradition so understood, the central point is that the problem of explaining cognitive phenomena will comprise two activities, at least analytically distinguishable: The decomposition of the cognitive process in terms of the necessary structures, subprocesses and operations that work in a coordinated way to perform a certain activity –what Bechtel and Richardson call structural and functional decomposition– and the subsequent localization of the specific brain structures responsible for these activities –what the authors understand in terms of a mapping between activities and operations, one the one hand, and the associated structural components, on the other.

This is where the central requirement that the resulting models reveal the causal mechanisms that produce or maintain the phenomenon of interest steps in. After the identification of a mechanism responsible for a given phenomenon, what follows is a description of its causal structure: That is, its component parts, its properties and activities as well as its organization. Kaplan and Craver (2011) call this desideratum the 3M constraint for explanation, insofar as there should necessarily be a mapping between model and mechanism in which the variables included in the model and the dependencies between them must correspond to the components, activities, and organizational features of the mechanism underlying the phenomenon at hand. An additional point of general agreement –with the clear exception of John Bickle's ruthless reductionism– has to do with the hierarchical organization of mechanisms: The parts of a mechanism can in turn be mechanisms and this opens the possibility of integration between different levels of description.

Again, what's interesting to the case at hand is that this scheme is so general that it encompasses a great diversity (1) at the different levels where problems are addressed –for example, in terms of both top-down (that is, taking the system's capabilities as starting point) and bottom-up (taking the system's structural components as starting point) approaches–, (2) regarding the privileged experimental, observation and modeling tools –from various neuroimaging techniques to methods centered on task analyses as so-called functional analyses (Cummins 1975) or indirect inferential methods (Glennan 2005) such as the use of reaction times– as well as (3) regarding the kinds of explanatory devices put forth –models of different kind and function, theoretical developments, mathematical apparatus and computational simulations. The decomposition of the system in terms of subprocesses and operations working in a coordinated manner, associated with their localization in the brain, is in this sense a strong pillar behind the interpretation of recent research in the field.

Now, I should anticipate a further point of agreement among many of the advocates of mechanistic ideas: Namely, the fact that these are generally seen as opposing a pluralistic stance regarding explanation, which allows for the possibility of other philosophical accounts of explanation applicable to cognitive science and neuroscience. In other words, arguments for the different mechanistic positions are usually accompanied by a thesis about their extensive applicability to large areas of cognitive science and specifically to neuroscience research centered on cognitive phenomena, and, even more so, about its comprehensiveness to account for the explanatory dimension that these fields exhibit. Later I will return to this point.

In the particular case of cognitive neuroscience, mechanistic positions are here divided between more or less reductionist positions, on the one hand, and integrationists, on the other. The main turning point concerns the relationship between molecular and cellular neuroscience, and psychology. In this regard the position developed by Bickle represents the reductionist end, while positions such as those of William Bechtel and Carl Craver represent moderate solutions that lean heavily on the concept of levels of description (for example, from microscopic levels of brain research to various kinds of neuronal systems), the idea of a hierarchy of mechanisms, and thus the possibility of a coexistence between different mechanistic explanations. This second kind of proposals has considered more thoroughly the problem of the dynamical approach and its explanatory contribution.

In recent years, it has been defended a number of opposing positions within this recent debate with an eye on the emergence of dynamical research in the contemporary panorama in the cognitive and brain sciences. These are positions facing the tension that arises between the novelty that this kind of approaches carries with it and the mechanistic model of explanation. One can distinguish three kinds of positions in the debate: Positions linking dynamical research to the introduction of a new kind of cognitive-scientific explanation independent of mechanicism, positions that frame this research within the classical covering law model, and positions that understand it as a variant of the mechanistic stance.

In the first case, dynamical models are understood as explanatory regardless of the fact that in many cases they do not take into account the underlying causal structures that lead to system-level dynamics (for example, Giunti 1997 and van Gelder 1998, for some early accounts, and more recently Chemero and Silberstein 2008, Stepp et al. 2011, and Ross 2015). Arguments here revolve around the descriptive and predictive advantages brought by the application of dynamical systems theory: "One key feature of such dynamical explanatory models is that they allow one to abstract away from causal mechanical and aggregate micro-details to predict the qualitative behavior of a class of similar systems" (Chemero and Silberstein 2008: 12).

Positions of this sort defend the idea of a certain novelty in the dynamical approach in terms of its descriptive and predictive potential. They also defend the idea that models that portray the systems of interest in such high level of description do not require further elaboration through decomposition and localization heuristics. Now, it should be acknowledged that there isn't here an anonymous definition of the explanatory nature inherent to these models. For example, recently, Ross (2015) turns to Batterman’s proposal of explanation through minimal models, and Gervais (2015), although more moderate regarding explanatory power, points to an additional factor brought by dynamical models inasmuch as they contribute to a greater understanding of the phenomena under study.

According to the second kind of positions, dynamical models offer explanations along the lines of hempelian covering law accounts (Kelso 1995; Walmsley 2008): That is, they show phenomena of interest as instances of a regular pattern. In particular, resorting to laws that govern the systems’ behavior is seen as central. It should be pointed out that the previously distinguished position recognizes in some cases certain proximity to the covering law model (see, for example, Stepp et al. 2011: 432); considering the case of systems neuroscience, Silberstein and Chemero (2013) also write of mathematical explanations approaching the hempelian model (see also Chemero 2009). However, I think we should distinguish both approaches because of the equivalence that, following Kaplan and Bechtel (2011), tends to be claimed between the covering law model and so-called “predictivism” –that is, the idea that the explanatory power of dynamical models strongly depends on its predictive ability–, a position that many advocates of the first position (such as the already mentioned Chemero, Silberstein, and Ross) avoid.

Finally, according to the third group of positions, dynamical models should be understood as descriptive tools that allow us to represent how certain complex mechanisms work and are in this sense part of the project of providing mechanistic explanations (Kaplan and Craver 2011; Zednik 2011; Kaplan and Bechtel 2011). This is perhaps the dominant position in the debate. The main idea is as follows: “There is no currently available and philosophically tenable sense of ‘explanation’ according to which [dynamical] models explain even when they fail to reveal the causal structures that produce, underlie, or maintain the explanandum phenomenon" (Kaplan and Craver 2011: 602). These models are accordingly understood as subsidiary to strategies to identify and characterize cognitive mechanisms. As this is the most developed and widely defended position in the debate, and as much of the following considerations will be aimed at assessing this preeminence, it is appropriate to look with some detail to a particular case. I will consider the proposal by Zednik (2011), which will be further examined later on.

Zednik's main thesis is that dynamical research is often used to describe cognitive mechanisms and thus constitutes a kind of mechanistic explanation. His strategy is to make explicit the principled compatibility between mechanisms and dynamical models and analyses (similarly, Gervais 2015 defends this possibility): Briefly, the differential equations, concepts, and graphical tools courtesy of dynamical systems theory constitute descriptive resources and in this sense can be interpreted as representations of cognitive mechanisms. To show this, Zednik takes two cases of successful research in developmental psychology and computational neuroethology (respectively, Thelen et al. 2011 and Beer 2003). He then strives to show in what way these models approximate the mechanistic strategy: In particular, according to the author, these are reductive and can invoke representations, a theoretical construct crucial to a standard type of approach in cognitive science broadly conceived.

Zednik, in a similar attitude to the work of Bechtel, doesn't take this as a critical point for the dynamical approach. He defends its novelty and the contribution it makes: In particular, he understands this kind of models as specially suited to describe extended (that is, beyond the narrowly conceived mind-brain construct) and complex (for example, cases of coupled dynamical systems where continuous reciprocal causation phenomena may arise) mechanisms. However, as can also be seen in the compatibilist strategy developed by Bechtel (1998, 2001, 2008), the underlying idea is that such models should be understood as descriptive tools particularly suited to address the evolution of certain complex mechanisms and thus are not only compatible with the mechanistic strategy but a part of it: In short, if these models explain it is because they describe mechanisms. Ultimately, as I already stated, the guiding principle of positions of this kind is the defense of certain exclusivity of mechanistic explanation in the cases of cognitive science and neuroscience, and, above all, the lack of a viable philosophical alternative, above and beyond the acknowledged novelty and benefits behind the adoption of dynamical systems theory.

3. Some cautionary considerations…

Here I propose a critical approach to the way the presented debate unfolded in recent years, along the lines of the three proposals just presented. In particular, I want to advance some general cautionary considerations on the debate's development, with some suggestions for its redirection. The following observations stem from my take on its genesis. I assume that the issues surrounding dynamical research constitute a very specific application of debates pertaining to general philosophy of science to the still young field of cognitive neuroscience: In this sense, I think that they inherit some problematic biases that deserve to be highlighted.

Koertge (1992), in an incisive review, acknowledges the prolific philosophical work on the issue of scientific explanation by the end of the last century, but also notes a striking lack of reflection on the nature of the problem and why it matters. She thus proposes to consider what kind of problems a philosophical theory of explanation should in fact solve. I think this question takes on another dimension when applied to the fields considered here, and I think it is very relevant to formulate it in this case.

This has to do, firstly, with taking seriously Koertge's recommendation for the case at hand and, also, with a healthy preliminary precaution prompted by the still incipient nature of the research areas at the center of the philosophical debate presented here: In general, it is necessary to define appropriate epistemological categories for this kind of applications to the study of the related explanatory practices, and to adapt other categories used in the literature with a focus on more established scientific areas.

In what follows I take into account two aspects of cognitive neuroscience that are very relevant to this: Its diversity and its stage of development. In particular, a central issue that I will later outline is that often the question of scientific explanation in cognitive science and neuroscience seems almost to be equated to the degree of success or inherent value of a particular model. This way of framing the issue of explanation is not consistent with what I see as the main distinctive features of the field. With an eye on these caveats, I will then try to show how looking at the problem from a context of inquiry interestingly redefines it, specially in the direction of rejecting questions, common in the literature, such as: Does this particular model explain or not?

Firstly, I set some limitations on the problem of explanation concerning a field as young and diverse as cognitive neuroscience is. Secondly, I analyze a specific example –Zednik's analysis of a known case of dynamical research–, to show that its mechanistic interpretation is forced, and that in addition it focuses on issues that do not highlight the main benefits there obtained. Thirdly, I recapitulate my cautionary proposal on the limitations of the subject of explanation to reveal the plot of epistemological aspects behind cognitive neuroscience, while suggesting some of the interesting contributions that a kind of philosophical work more sensitive to scientific practice can yield.

3.1. A panoramic view on current cognitive neuroscience

To a first approximation, it seems safe to encourage some caution about the often assumed idea that there is a kind of property or relationship (causal relevance, unification, covering law, etc.) that captures the concept of explanation along various local applications. Again, Koertge says:

> It might be useful to classify scientific queries according to their logic and epistemological structure. It remains to be seen, however, whether the questions we would ordinarily consider to be requests for explanation fall within a single cell of the problem typology. (Koertge 1992: 97)

While the idea here is to lay a veil of caution on the issue of scientific explanation in itself, this attitude will be further reinforced and justified once we consider our scientific field of interest. Let's turn to the case of the dynamical approach.

The first attempts to define dynamical research's explanatory status were somewhat general in their treatment of the lines of research taken into account. In particular, adding to the widespread application of dynamical systems theory, the debate focused on cognitive science taken on a global basis. Let's take as an early example, the vigorous defense put forth by Tim van Gelder (1998) of what he called the "dynamical hypothesis" –centered on the idea that natural cognitive systems are a special kind of dynamical systems–, intended to effortlessly condense different research programs within the dynamical approach. In the first place, a hypothesis of this kind is very vague and nonspecific, and in this respect not very useful or informative, if it is not interpreted from a methodological perspective in terms of the characteristic mathematical and graphical resources that fuel a particular approach to cognitive phenomena within a defined disciplinary context. Secondly, it tends to promote the formulation of philosophical issues on the relevant scientific work which are strictly irrelevant from a philosophy of science point of view, such as is the problem of the underlying conceptions of cognition behind that work (see van Gelder 1997: 438-440).

The debate notably improved since its inception in the nineties, both on account of its sophistication and its closer look to particular scientific fields. The increasing use of case studies to illustrate the philosophical points is a clear sign in this direction. It is however necessary to favor another qualitative leap towards a debate more focused on the workings of contemporary cognitive neuroscience. A telling fact is how many of the main actors in the debate defend their positions, sometimes centered on some variant of contemporary basic neuroscience (Craver 2008; Ross 2015), sometimes on cognitive neuroscience (Bechtel 2008; Abrahamsen and Bechtel 2012), and sometimes on cognitive science in general (van Leeuwen, 2005; Chemero and Silberstein, 2008). Thus, often the philosophical debate doesn't seem to be standing on a homogeneous platform when it comes to the kind of research involved. These differences of focus in the philosophical work are crucial and should be clearly stated (see in this regard Revonsuo 2001)[1]. At the same time, they reveal the enormous heterogeneity that cognitive neuroscience provides for an eventual definition of its explanatory status.

The issue of heterogeneity in cognitive science has been increasingly tackled from a philosophical standpoint: We can mention Abrahamsen and Bechtel (2006), Dale (2008), Dale, Dietrich, and Chemero (2009), and Stepp et al. (2011), among other efforts that underpin this recent pluralist trend. Vis-à-vis the question of explanation, it is necessary to further highlight this point, though I think from a radically different perspective. In the mentioned cases, in fact, the issue is sometimes posed in terms of approaches and sometimes in terms of co-existing theories, that is, in line with the philosophical tradition that mainly distinguished between computational, connectionist and dynamical theories. Heterogeneity is thus seen as most properly theoretical: That is, defined by the idea of a cognitive system as a rule-following computational system that operates on symbols, or as a system that manipulates sub-symbols according to implicit statistical rules, or else, finally, as a dynamical system in ongoing interaction with its environment. Yet I believe that these categories are still too coarse to capture the heterogeneity of fields such as cognitive neuroscience. Here, another kind of diversity persists, that has to do both with the phenomenon under study (in standard terms, the explanandum) and the explanatory devices deployed to study it: A kind of diversity much more eloquent to the problem of explanation.

In the case of neuroscience, this diversity has been recognized both by working neuroscientists and philosophers (for example, Sullivan 2009; Bickle and Hardcastle 2012: 5; Author 2015). Now, in general, debates about the diversity of practices and styles of research are rarely held from a philosophical point of view; what prevails instead is a sort of tacit agreement on the extent to which this diversity doesn't substantially affect the problem of explanation. Regarding the kinds of phenomena under study, for example, there is a tremendous variability among the ranges of accuracy and spatiotemporal scales at which they are described. As is known, in cognitive neuroscience, there are ongoing research programs addressing the overall activity of the whole brain, both at experimental (for example, Rodriguez et al. 1999) and modeling levels (for example, Deco et al. 2015), as well as lines of work centered on the impact that genetic regulatory activity in the neuron has on psychological effects (for example, Ebbinghaus spacing effect in long-term memory, discussed by Bickle and Hardcastle 2012).

Here I think it is convenient to call attention to the lack of clarity in the notion of "levels", often used in the philosophical literature to capture this aspect of the aforementioned diversity. In the case of dynamical research, a common strategy is to restrict its explanatory power to systems or high-level neuroscience, while recognizing the success of mechanistic models in research on genetic factors, neurotransmitters, or the action potential (for example, in line with Chemero and Silberstein 2008, and Silberstein and Chemero 2013). This is another way to define pluralism in the sense of a peaceful coexistence between models directed toward different levels of the brain's structure and which provide different explanatory benefits.

Now, in cognitive neuroscience it is often very difficult to define at what "level" a specific data set belongs to since in general the development of an experiment in order to answer a given question involves assumptions spanning different ranges and degrees of accuracy (Woodward 2015, draws attention to this fact, despite making extensive use of the concept). This can be clearly seen attending to the growing tendency of the joint study of a particular problem or phenomenon based on different imaging and brain probing techniques: The point is that these techniques offer some kind of access to limited spectrums of spatial and temporal accuracy as well as very different aspects of brain activity (see in this regard the case of combined fMRI / EEG studies, such as Rosa et al. 2010 and Huster et al. 2012).

Another more general aspect of this is reflected in the typical experiment performed on human subjects in cognitive neuroscience, where the researcher inevitably has to define assumptions both regarding the cognitive or psychological phenomenon of interest, and how it comes into play in the context of a task, as well as regarding the neural system that is assumed to be involved. In these cases, we are left with a situation where it is not entirely clear if what we can call the explanans level is strictly on a par with, or contained within, the explanandum level. It is also worth mentioning the extensive field of work devoted to highly diverse animal models, which greatly diversifies and extends the spectrum of phenomena under study.

Regarding the adopted kind of experimental apparatus, there is a wide variety of different techniques to explore brain's activity and, specially since the introduction of functional imaging methods in the nineties, the choice of a particular technique carries with it a set of assumptions with different degrees of generality and at different stages of research (population sampling, experimental design, data processing and analytical approach, data interpretation, and so on). Given the novelty of these techniques, which are in fact largely responsible for the development and unprecedented growth of research in cognitive neuroscience, it wouldn't be wrong to view them as a field in which we are still beginning to learn on the limitations of the instruments adopted in experimental contexts. This is particularly so for the purposes of philosophical argumentation. I later expand on this point.

We must also consider computational and simulation approaches to neural systems, in which a considerably volatile variable is the degree of biological detail that is incorporated into the model or, in other words, the measure of interpretability in neuroanatomical and neurophysiological terms: Consider, for example, the distance separating two widely known models such as Hodgkin and Huxley's model of the action potential and Hopfield's model of associative memory. This rich line of work involves heuristics and theoretical goals sometimes very far away from the experimental work already considered (as will be clear in the example discussed in detail below).

All this is to suggest that the situation of diversity and heterogeneity typical of cognitive neuroscience cannot but affect theorizing on its explanatory status. In other words, the idea of "pluralism" and its relevance in this regard should be taken carefully, installed as it often is in a tension between frameworks, approaches or theories, so as not to risk taking an easy and not very telling solution about the complex research practices on the market. Again, concerning the debate on explanation in the field, a major problem when we consider these warnings is the selection of scientific examples used to substantiate the philosophical positions presented above.

In many cases, mechanistic positions are argued for on the basis of examples taken from very successful tales of very elaborate, basic neuroscience models: A case again and again revisited is that of studies stemming from the already mentioned Hodgkin and Huxley's original model of the action potential's dynamics, which managed to detail the activity of potassium and sodium channels that causes the electrical changes in the neuron's membrane conductance. Kaplan and Craver (2011), for example, use this example to reinforce their mechanistic position, and then claim:

> Cognitive and systems neuroscientists have yet to make the kinds of inroads in their domain that Hodgkin, Huxley, Hille, and others made in understanding the electrical properties of neuronal membranes. Given that one expects cognitive mechanisms ultimately to be composed of lower-level mechanisms of this sort, in a manner that might be illustrated in a telescoping hierarchy of mechanisms and their components, it would be most tidy and parsimonious if the ideal of mechanistic explanation were to be extended from top to bottom across all fields in neuroscience. (Kaplan and Craver 2011: 611)

The scope and impact of this assumption is what I find problematic. On the one hand, it rests in turn on an at least questionable homogeneity assumption for neuroscience. It is to be expected that strategies and research heuristics as well as the kinds of explanatory apparatus needed to study the activity of potassium channels in action potential generation would not be the same as those that will eventually be successful in explaining the most varied and evolutionarily recent cognitive phenomena.

Moreover, the authors' assumption hides a normative assessment of the current path of neuroscience and, in particular, on its uniform continuity in the direction of addressing increasingly more complex phenomena. Even though many proponents of versions of the mechanistic position are careful not to limit this kind of explanations to bottom-up approaches, here we can see a bet on a progressive "no surprises" advancement from basic towards more complex phenomena, at least in terms of the factors involved in the explanandum, all eventually sheltered under the same model of explanation. At a minimum, the risk is neglecting the relative independence between different models in terms of their scope and / or neurobiological detail. This, again, is a seldom raised questionable point.

It is noteworthy how this wakeup call has a considerable affinity with the idea developed by Matthewson (2011) of a target-oriented approach for models' assessment, and in particular of the balance achieved between their many and often contrasting desiderata. Although the author's reflections point towards population biology (in line with the pioneering ideas of Richard Levins), I think that weighing the class of modeled system vis-à-vis the kind of epistemic benefits extracted from the modeling effort is very eloquent for the field of interest, and particularly considering the significant heterogeneity of target phenomena. An in-depth analysis of this point for cognitive neuroscience, although of great interest, exceeds the goals of this paper.

It is also worth noting that this caveat takes on special strength bearing in mind the implementation of nonlinear systems mathematics. This is so to the extent that, as I noted earlier, it is not restricted either in principle nor in practice to any defined area of neuroscience or cognitive science research. The extensive applicability of dynamical systems theory, which can be appreciated considering the traditional conception of cognitive science as an interdisciplinary domain as well as in the most diverse areas of neuroscience, should be interpreted as a serious hurdle for a definition of the explanatory status that comes with it.

The problem of case study selection is closely connected with the problem of cognitive neuroscience's state of the art, which also strongly influences the issue of scientific explanation in the field. I already referred to the novelty of many of the most widely used experimental techniques. In this sense, the large number of studies dedicated to evaluate the stability of a particular technique, the robustness of the results it generates, the strength of the various assumptions on which it is based, the factors affecting the statistical power of the results, the development of analytical approaches for the data, and so on, is remarkable. In short, the methodological aspect of experimental approaches breeds a host of healthy and widespread theoretical concerns. This can also be noticed in the interplay between experimental research and the multifarious contributions of computational simulations.

If we now turn back on the application of the mathematical tools of dynamical systems theory, it would be appropriate to think of it as an attempt to recover common and more established methods within physical science, an attempt that initially characterized the pioneering cybernetics movement. In this sense, a good way to look at it is as an attempt to calibrate a language and a set of analytical and graphical tools that have shown a great expressive power in various areas of science, in order to apply them to a different area of scientific exploration. This is another aspect of the novelty and the associated stage of development of the field of interest, particularly with regard to modeling practices[2].

Taking into account these considerations, we can understand why, even from the dynamicist front of philosophical debate, positions have hardened. A recent example is Silberstein and Chemero (2013), who defend the presence of genuine explanations with examples taken from systems neuroscience, particularly simulations that adopt graph theory and dynamical systems theory. It is by no means a minor point that the lines of research and the results there obtained, selected to defend this idea, are all cases of computer simulations: In particular, studies (specially from research carried out by Olaf Sporns) of the topological properties of the architecture of simulated neural networks. In no way am I suggesting that these results are minor or irrelevant, but clearly they only represent one strand of a very exploratory approach to the activity of biological neural networks.

Now, dealing as we are with a very recent development in the research style in cognitive science and in particular in its neuroscientific variants, it seems appropriate to try to avoid the tendency to pigeonhole specific results as delivering particular kinds of explanation. Accordingly, it may result to be premature and even counterproductive to associate research based on dynamical systems tools with a particular philosophical account of explanation, such as the hempelian (Walmsley 2008) or mechanistic (Zednik 2011) model, and this is precisely what the positions presented above do (respectively, the second and third position).

In the specific case of the covering law model, it strikes as too strict a requirement to necessarily appeal to laws at some point in the explanatory effort: This requirement is not inherent to model building via systems of equations nor to their analysis through dynamical systems theory, a contention agreed by both mechanicists and advocates of the first position previously presented. Setting aside the many problems posed against the hempelian model (see in this regard Kaplan and Craver 2011), I take its general application to cases of dynamical research as a risky and, in particular, unnecessary leap aimed at defining its style of work and contribution in terms of explanatory status.

As I mentioned, the cases of Thelen et al. (2001) in developmental psychology and Beer (2003) in computational neuroethology are taken by Zednik (2011) to show that, even in cases of clearly dynamical research, we can talk about mechanisms as far as its explanatory contribution goes. Although we are dealing with very simplified models at a neurobiological level, spanning a cognitive agent's brain, body and environment, Zednik's effort shows in what sense a mechanistic interpretation of dynamical models is sometimes downright impoverished and falls into a premature definition of the problem of explanation. It is thus important to take this kind of arguments into thorough consideration.

3.2. Case study: The dynamics of adaptive behavior

As I anticipated, Zednik's main point is that "the differential equations and graphical representations that figure in many dynamical explanations can be, in principle as well as in practice, interpreted as representations of cognitive mechanisms" (Zednik 2011: 247). To show this, he relies on two case studies from dynamical research, which he contrasts with the well-known Haken, Kelso and Bunz model of bimanual coordination. One is the model of the A-not-B error in developmental psychology, carried out by Esther Thelen and collaborators. The other, on which I will focus, is a model of categorical perception developed and analyzed by Randall Beer.

The underlying idea the author argues for is that, in both cases, the strategy of (functional and structural) decomposition typical of mechanistic research continues to operate as a central heuristic and that, in turn, the path for an eventual localization of the identified parts and processes is left open: That is, the idea that, when it comes to the explanatory contribution of both models, this is of an essentially mechanistic nature. In order to follow my analysis of the explanatory character of the model of categorical perception and how Zednik fails in its interpretation, it is necessary to briefly look into the research program carried out by Beer.

Beer's research is distinctive because it lies at the periphery of cognitive neuroscience. The model we'll focus on (Beer 1996, 2003) belongs to the program of computational neuroethology, centered on the simulation of basic cognitive phenomena from simple idealized models of complete cognitive systems, that is, comprising aspects of the brain or control system, the body of a cognitive agent, as well as the environment in which it operates. In particular, it addresses what Beer calls "minimally cognitive behavior" (Beer 1996), whose additional interest lies, on the one hand, in its novel character as a research program and, on the other hand, in its explicit concern for the cognitive nature of the phenomena taken into account. Examples of such behavior are object discrimination, categorization and manipulation, short-term memory and selective attention (both explored in Slocum, Downey, and Beer 2000), associative learning, the ability to signal other objects or agents, etc. The model we are interested in here is Beer's first exploration in "minimal cognition": It is the model of an agent capable of categorical perception, originally developed in Beer (1996) and analyzed in Beer (2003). This more recent analysis is the relevant result for our interests, and the one which Zednik focuses on.

In the original computational model, Beer (1996) used genetic algorithms to evolve continuous-time recurrent neural networks for the control of the behavior of model agents in a specific environment. Beer's evolved agents have the task to catch circular objects and to avoid diamond-shaped objects. The kind of categorical perception (for example, Harnad 1987) at stake here is the ability to selectively perceive differently shaped objects (circles and diamonds), that is, to distinguish based on the continuous information received by sensory organs different categories among the different objects presented to the simulated agent. Beer agents “live” in a two-dimensional world and only move along a horizontal line at its base. From the top of the environment, diamonds and circles fall down until reaching the base where the agents are. Once the agents complete their development stage, their task is to catch circles and avoid diamonds: In order to do this, they have to detect, through their sensors, the objects’ geometrical shape, their relative distance and rate of approximation, and thus take some suitable course of action, through their motor neurons (displacing the agent to the right or to the left).

Beer (2003) analyzed several aspects of the behavior of the most successful simulated agents. Once the results of the simulations had been obtained, three stages of Beer’s work can be distinguished: A descriptive stage of the resulting behavior, an exploratory stage where more precise information is extracted from the simulations, and a properly analytical stage where abstract features of their dynamics are examined. The first stage focused on the agent’s behavior: For example, different strategies employed are distinguished and task performance is assessed. In the second stage, Beer is concerned with analyzing what he calls “psychophysical experiments”, in which he manages to determine specific aspects of the agent’s categorical perception by altering certain crucial parameters, such as gradually varying the objects’ shape or abruptly shifting the object’s identity at different times of its fall. Through this manipulation, he was able to identify the presence of subroutines in successful cases of categorical perception, the crucial object features enabling the discrimination between diamonds and circles, and the time dependence of the way the agent reaches its “decision”.

The third and final stage is devoted to analysis of the simulations through the language of dynamical systems theory, in order to examine and rigorously characterize various features of the system instantiated in the model as well as to generate accurate predictions about possible agent's behaviors. Here Beer for example appeals to the shape of the trajectories that describe aspects of the agent's behavior or the presence of limit cycles, equilibrium points, discontinuities, bifurcations, and so on. This is the most relevant analytical effort in terms of examining the model's dynamics. These dynamics are addressed at three levels: At the global agent-environment system level, at the more specific level of the interactions between agent and environment, and finally at the neuronal implementation level of the active agent (which would thus represent the more properly neuroethological contribution).

In his assessment of Beer's work, Zednik understands that the main result there obtained lies in the decomposition of a mechanism responsible for a complex activity such as categorical perception into its component parts, their operations, and their organization as a mechanism. While Zednik would agree that the model's contribution rests largely on the third analytical stage, based on the application of the mathematical tools of dynamical systems theory, he takes this to be so on account that this application allows to describe the activity of two components of a global system: The embodied brain, on one side, and the environment, on the other. The dynamical analysis allows for a precise description of the operations of both components and the way they influence each other.

So, according to the author, "Beer describes the component parts, component operations, and organization of a mechanism for perceptual categorization via active scanning" (Zednik 2011: 254). To Zednik, turning to the level of neural organization, Beer even describes a further neuronal mechanism, centered on the contribution of some neurons in the intermediate layer of the neural network, thus exhibiting the hierarchical aspect defended by many mechanistic philosophers. The author goes so far as arguing that Beer's psychophysical experiments –in the second stage of the research– meet the requirement made by Craver that an adequate description of a mechanism would in principle make possible its manipulation and control (cf., Zednik 2011: 259-260).

I believe that this interpretation of the contribution due to Beer's dynamical analysis is extremely simplistic and impoverished. Firstly, this is because the significance of the different analytical efforts by Beer in the context of his line of research is not adequately assessed. The case we are considering is different from the one I commented regarding the position defended by Silberstein and Chemero (2013). Although this too is a simulation that operates far from specific neurobiological details, we are facing a particular kind of approach whose basic aim is to weigh the scope and implications of certain key concepts. The resulting models are analogous to what Mitchell (2009: 211) calls "idea models", that she connects with the recent trend of complexity sciences.

In this case, consider that the author's intention is explicitly "not to propose a serious model of categorical perception, but rather to use this model agent to explore the implications of dynamical explanation for cognitive agents" (Beer 2003: 210). Thus, this kind of models is very eloquent to inspect the set of theoretical projections and assumptions the researcher starts with, concerning the delimited object of study. This is due to the great freedom it affords her: A freedom to select parameters and to alter them in accordance with the specific modeling questions that may arise in the course of research. It is within this line of work so understood that the contribution of Beer's simulations should be assessed.

Furthermore, it seems misguided to argue that Beer actually describes parts, operations, and organization of some assumed mechanism, following, as it were, a mechanistic explanation. I think he does not, and that this maneuver by Zednik brings us dangerously close to a sort of dull "panmechanicism". Firstly, the "parts" considered by our philosopher, on which Beer spends much of his dynamical analysis, are the embodied brain and the agent's environment: These do not in the least seem to be parts whose coordinated operations could be considered the activity of a mechanism, not even according to the weakest mechanistic accounts. This at least applies to the first two stages in the study of the simulations and to the first two steps of the third analytical stage, all focused on the agent's behavior dynamics and on its interaction with the environment.

Now, let's consider the situation regarding the final effort of the third analytical stage, focused on the neuronal implementation of the simulated agent. Zednik understands this as the description of a "*neural-level* mechanism" (cf., Zednik 2011: 254). Here an epistemological precision internal to the logic of Beer's computational neuroethology can be enlightening: As he explicitly states (cf., Beer 1997: 271), he dos not assume any commitment to the architecture of his agents (that is, continuous-time recurrent neural networks); this means that behind the adoption of this particular architecture there isn't an additional hypothesis about what would be the appropriate neural mechanisms behind, for instance, categorical perception. It is important to emphasize this as it clearly reveals the exploratory nature of Beer's research as well as the abstract or conceptual domain of the addressed problems. The selected kind of neural network turns to be useful for certain theoretical objectives, which could be tackled through other means.

The exploratory character of Beer's research is a third aspect where Zednik's interpretation is misguided. As I said, in cases such as this where there aren't strong constraints from, for example, experimental data on a particular cognitive phenomenon or neuroanatomical details about a given species, the researcher's control on the model is very high. This is noticeable in our case both in the elaboration of the models and simulations, and in their analysis. Consider for instance the psychophysical experiments in the second stage of our revised model's analysis: There, Beer freely alters certain key model parameters simply to see how the resulting changes affect the simulated agent's categorical perception performance.

Unlike a focus on decomposition, the focus here is instead set on the rigorous characterization of the dynamics of a previously defined system, dynamics that can occur both internally to the brain or control system and in its interaction with aspects of the agent's body and, in turn, in its interaction with the environment. Accordingly, an adequate perspective on Beer's work is to identify the phenomenal target of his research with the dynamics of adaptive behavior, from different fronts, rather than with the neural systems responsible for a given cognitive skill: This focus can be equally defined in terms of the basis of implementation in some sense responsible for the behavior (namely, neuronal dynamics), in terms of the search for general behavioral principles, or even in terms of more properly cognitive aspects eventually involved (standardly understood in terms of information processing).

Much of the fruitfulness of Beer's model lies in the third analytical stage, tied to the language of dynamical systems theory. This third stage shows Beer's object of study as an abstract domain that we can associate, in this particular case centered on minimally cognitive behavior, with some elementary notion of cognition. The extent to which this confined search space is highly abstract can be appreciated if we turn to the methodological decisions at this stage of Beer's work. From a methodological perspective, what's remarkable for the establishment of a fertile research program is the conceptual precision obtained through analysis courtesy of dynamical systems theory.

As we saw, Beer undertakes this analysis on three levels at which different aspects of the system are taken into account. The relevant point here is that the global analysis of the brain-body-environment system, the analysis of agent-environment interaction, and the analysis of the active agent's neuronal implementation all study abstract features of the simulations' dynamics. The analysis' exhaustiveness and accuracy are the epistemic values or contributions that turn the analytical effort informative[3]. While the second is courtesy of the advanced development of nonlinear dynamical systems theory, the first is noticeable, for example, in Beer's treatment of the complete state space of the system as it is defined in each case, that is, according to the variables and parameters taken into account at each analytical level.

These considerations stemming from our case study aim, on the one hand, to strengthen the idea that the philosophical analysis of a specific model in terms of its explanatory contribution should not be made independently of the line of work or research framework which contains it: These provide a situated context that often defines evaluation criteria and research goals. On the other hand, they aim at discouraging a hermetic kind of take on scientific products, that tends to define their explanatory character, in a field that is very diverse, dynamic, and, above all, still in a germinal stage of development. The case considered here is only a particular sample of the situation I described earlier for the general field of cognitive neuroscience.

3.3. Critical thoughts on the philosophy of explanation

A philosopher who has recently defended ideas in tune with those here espoused is Woodward (2015). The general thesis he objects is the contention, often defended in the literature and already criticized here, about the exhaustive applicability of mechanicism to all areas of neuroscience and cognitive science. Woodward, on the one hand, takes a cautious attitude towards results obtained in neuroscience, both given its state of the art and, above all, given the bounded, situated targets set by experimenters and modelers, a point I tried to illustrate through Beer's case study[4]. The author states:

> The picture just sketched [...] is thus opposed to an alternative picture according to which a theory that explains any explanandum satisfactorily must be a "theory of everything" that explains all aspects of the behavior of the system of interest, whatever the scale or level at which this is exhibited. (Woodward 2015: 12)

Moreover, Woodward sees in mechanistic accounts an attempt to provide definitive criteria for scientific explanation. In particular, the 3M constraint, of a mapping between model and mechanism, is proposed, in the version of Kaplan and Craver (2012), as a general criterion which applies regardless of the kind of phenomenon and the kind of interests that motivate modeling, and which accordingly can be used to sanction the explanatory value of individual models. The problem is that this sort of general criterion is sometimes not available in scientific practice: "Models can successfully convey dependency information in surprisingly indirect ways that do not require this sort of mirroring or correspondence of individual elements in the model to elements in the world" (Woodward 2015: 21).

This is a way of applying the caution that, as we saw, Koertge recommended for general philosophy of explanation to our field of interest. An attitude of this kind, as that developed here, is one that undermines the "fatalistic" stance according to which we should be able to choose between two different models of the same phenomenon according to their explanatory credentials. This is sometimes simply unfeasible. Returning to the debate on the role of dynamical models, we can see how other negative consequences of views of this kind arise.

The counterpart of the idea of a comprehensive and definitive criterion for neuroscience is the recognition of the lack of other possible explanatory criteria. As I noted above, Kaplan and Bechtel (2011) reject the dichotomy between mechanistic and dynamical explanations in the sense that in cognitive neuroscience an alternative to the explanations of the first kind is simply not available[5]. Now, the next step in this maneuver is to defend the compatibility of dynamical models with the more common mechanistic strategy: That is, to show what contribution these models make within a comprehensive mechanistic logic.

However, I understand that the increasingly invoked requirement (Bechtel 2001; Kaplan and Craver 2011; Kaplan and Bechtel 2011; Zednik 2011; Abrahamsen and Bechtel 2012) of a convergence towards mechanistic approaches –for example, in the previously discussed case, to the extent that one can contribute to the heuristic of decomposition and localization– unnecessarily anticipates and restricts future dynamics of theoretical change in the direction of the quest for mechanisms. For example, Kaplan and Craver (cf., 2011: 618) take the very fact that Kelso and colleagues ventured into the neural basis of their model of bimanual coordination as an expected step in the direction of the description of mechanisms. This not only ignores the specificity of Kelso's neuroscientific research, based on the search for large-scale spatiotemporal patterns (for example, Engel et al. 2010), but also conflates plain neuroscience research with the search for mechanisms, again bringing us close to vacuous theses. I follow in this sense Dupré’s warning (cf., Dupré 2013: 28) about the presence of a “serious danger of vacuity” in many claims by mechanistic philosophers.

Moreover, one would think that, to the extent that such restrictions urging a reconciliation with mechanistic heuristics are mounted on reductionist methodological requirements, they fail to account for the role of dynamical models in many areas where they have been successfully applied (as, again, developmental psychology as exemplified by Thelen et al. 2001, work on motor coordination by Kelso and colleagues, or the line of research exemplified by Beer’s work, according to my interpretation). In this light, a perspective that fails to contemplate the most interesting contributions of dynamical research –understood as it is as subsidiary to strategies aimed at identifying and characterizing cognitive mechanisms and their structural bases– comes into view. In this sense, despite my criticism of Silberstein and Chemero (2013) about the provisional nature of much of dynamical research, I agree with the general idea they defend, according to which just neglecting the possibility of other kinds of non-mechanistic explanatory contributions in neuroscience and cognitive science constitutes an unnecessary step.

More broadly, a final cautionary consideration is a criticism about the dominant role that debates about the models’ explanatory profile often occupy at the expense of other epistemic or cognitive goals they may have. This point has ultimately to do with an overall appraisal of recent developments in the philosophy of cognitive neuroscience, largely dominated by the problem of explanation: This problem has clearly been a very important factor of growth in the area, but today I think it has negative consequences for its ensuing development. As we saw, this is so owing to the extent to which it delimits the set of relevant issues and it biases other issues that are addressed from the viewpoint of this overriding problem.

As an example, scientific integration is worth noting, as a set of practices within contemporary cognitive neuroscience, which has assumed a life on its own for the last two decades and a strong impetus in recent years. One could say that it is a recent trend (Stewart and Walsh 2006; Gazzaniga, Doron and Funk 2009; Aminoff et al. 2009; Cooper and Shallice 2010; among others) facing the persistent dispersion of areas and approaches. Now, the few philosophical accounts devoted to this concept are largely installed on the problem of scientific explanation, where the question of reductionism is fostered, and on problems framed in terms of psychological and neural levels of research. An example of the first kind is Craver (2005) and an example of the second kind is Bechtel (2002). (I try to counter this situation in Author 2015.)

Another counterproductive aspect of this dominant role explanation enjoys is that often the explanatory character of a particular model is taken as a critical and sometimes exclusive epistemic criterion for its evaluation: A certainly restrictive view on certain aspects of the model. This is exactly what we saw in the case of Beer's simulations, narrowly conceived as part of a decomposition-and-localization approach. In a quick glance at dynamical research, the ability to raise new questions is for instance noteworthy. This is an apparent point in Beer's research but that has also recently been portrayed by Ross (2015) through the canonical model of neuronal excitability: In this case, the research aimed at providing a model of class I excitability, originally described by Hodgkin, that could take into account the overall behavior of neurons within a very diverse range of their molecular detail. Here much of the novel contribution lays in the original formulation of the problem.

The same can be seen in the case taken by Silberstein and Chemero (2013). The local objective of many computer simulation studies is to make out aspects of neural systems that are worthy to be explored experimentally. This is the case of the branch of systems neuroscience, based on connectivity studies, considered by the authors. However, assessing the identification of topological features in simulated systems of large-scale distributed neural networks in strictly explanatory terms is not doing justice to the kind of approach at hand, or at least it amounts to limiting its scope. Another aspect in which research based on nonlinear dynamics techniques has plenty of room for development is the study of the temporal dimension of neural and cognitive systems: The ability to formulate empirical questions could be appraised, for example, in terms of the extent to which such research offers new insights into the role of temporal structure in behavior and cognition.

In general, a view of scientific research framed more in terms of specific problem solving or in terms of answers to particular questions reveals that only some of the questions that stimulate scientific research concerns explanation, while it would allow to highlight more features of models. In fact, the already mentioned example taken by Ross can also be seen in terms of unification, that is, to provide a comprehensive model of a disparate set of phenomena. But we could also take the role in generating hypotheses, mere predictive capacity, the ability to support counterfactuals, control over whether and how a certain phenomenon occurs, the description of complex patterns of behavior, identifying formative principles behind these patterns, the ability to deal with the typical instability of neural systems, exploring key concepts (in the case of idea models), the ability to define the robustness of an experimental technique, the role of guiding experimental search, or the heuristic role in developing experimental designs, all as common standards or criteria in cognitive neuroscience research, that do not necessarily result in greater explanatory power and are by no means limited to it.

These are all aspects that tend to be obscured by the tendency to study and analyze models only as finished products aimed at explaining and not as part of a larger process in neuroscience research. The opposite trend that I favor here would involve an approach that takes a given model from the standpoint of its development and the actual reasons behind it, and, at the same time, a target-oriented approach in the sense of the already mentioned Matthewson (2011). It is also worth noting that dynamical research, only exemplified here through Beer's case study, particularly illustrates the multiplicity and diversity of models' epistemic virtues, unlike the idea of a predominant feature over others (namely, the description of mechanisms).

It should finally be noted that this tendency to bias dynamical research (or any other set of research programs!) towards the quest for explanations enables the disqualifying tenor of certain positions in the debate: It is in this sense assumed that model evaluation is provided exclusively in terms of explanatory credentials. Here I am thinking of positions that philosophers such as Bechtel have been maintaining for a while: "Seeking dynamic models in the absence of a program of decomposition and localization [...] may produce vacuous science" (Bechtel 2001: 498). This is not a minor point as it reveals a certain disregard of, for example, prediction and the description of global patterns of behavior as, at least, subsidiaries to the main, purely explanatory, goal. Almost all actors in the debate show the remarked bias: Specially, those who give prominence back to the dynamicist contribution from the mechanistic horizon (as do Zednik, Bechtel, and Craver, among others).

At the same time, this biased perspective on scientific explanation fuels the reactive tendency, from the opposite front of debate, to claim legitimacy for the models that do not fit in. An example of this is the paper "Philosophy for the rest of cognitive science" by Stepp et al. (2011). As we saw, the concern here is to defend the explanatory nature of certain dynamical models (Stepp et al. 2011: 432-433) inasmuch as they are similar to covering law explanations and support counterfactuals. This is also a byproduct of the assumed equivalence between the epistemic value of a model and its explanatory status. Crucially, this also shows how the role of much of the philosophical debate vis-à-vis the related scientific disciplines is clearly inserted in a context of justification. Taking up once again Koertge's recommendation, the need to question the specific nature of the problem of explanation in a field such as cognitive neuroscience strongly imposes itself back again.

4. Conclusion

Facing the current rise of the "new mechanistic philosophy", I took into consideration the query raised by Koertge (1992) about what kind of questions should a philosophical theory of explanation answer, now in the context of recent research in the cognitive and brain sciences. I thus tried to outline some undesirable consequences of the prevalence of mechanistic philosophy of explanation in the field, focusing in particular on the dispute over how dynamical models fit in.

On the one hand, I attended to the recent debate on cognitive-scientific explanation as an adaptation of a classic debate in general philosophy of science that in this sense carries with it a universalist assumption of a single model of explanation for every research endeavor. While originally mechanistic positions were mainly proposed to account for the explanatory practices in molecular and cellular biology, I argued that their application to the field of interest is an unwarranted step. I suggested that the revised debate about the explanatory status of dynamical research is parasitic on a hermetic vision of cognitive scientific endeavors as well as on that unitarist assumption about a privileged model of explanation applicable to different research styles. To this end, I focused on a characterization of the diversity and state of development of contemporary cognitive neuroscience.

On the other hand, I argued that this debate obscures important distinctions concerning contemporary neuroscientific practices and leads to inappropriate conclusions regarding the interpretation of the results there obtained. I thus turned to a computational neuroethology analysis offered by Beer and its mechanical interpretation carried out by Zednik (2011). In line with authors such as Silberstein and Chemero (2013) and Woodward (2015), I criticized the idea that mechanistic models are the only kind of explanation in neuroscience and cognitive science, while at the same time discrediting the problem of defining a class of explanations for the case of the dynamical approach as an attempt to scrutinize its potential and effective contribution. I argued instead that philosophical models of explanation should be more sensitive to disciplinary differences and the associated contextual factors, and that at the same time they should rather ponder the relation of explanation to other epistemic goals of scientific research: A trend of this sort takes on particular interest considering the extraordinary number of theoretical and methodological assumptions underlying any individual line of work, both considering the kind of phenomenon under study and the means deployed to study it.

Finally, I advanced a more general criticism regarding the overarching role of explanation to the detriment of other epistemic goals and the way that in recent years it has concentrated philosophical attention on the scientific field of interest. In particular, I pointed out the danger of reducing dynamical approaches to a quest for explanations. I favored, as a projection for future elaborations, a more procedural look at modeling practices in cognitive neuroscience to the extent that this kind of view would display further aspects of the models as well as a more varied spectrum of the roles they play, over and above the search for cognitive and neural mechanisms. Although it was not the focus of my efforts here, I think philosophical discussion could greatly benefit from a more careful attention to some of these aspects of modeling and their role in cognitive neuroscience.

Notes

1. It is worth mentioning that I do not endorse the position Revonsuo (2001: 57-58) presents regarding dynamicism, which he sees as a recent version of functionalism in cognitive science; however, I highlight his warning on the importance of directing our reflections to more defined scientific fields.

2. It is interesting to note how Bickle (2006) appeals to what he sees as an unfitting assessment of neuroscience state of the art (in this case, molecular and cellular neuroscience) facing cognitive phenomena, as part of the motives behind the poor adherence to his radical reductionism in philosophy of neuroscience. Clearly, selecting the neuroscientific area that is object of philosophical reflection and subsequent evaluation has a major impact on the epistemological conclusions finally obtained.

3. Although other aspects can be highlighted: For example, in a philosophical take on this same model by Beer, Chemero (cf., 2009: 38) highlights its predictive benefits and its ability to support counterfactuals.

4. Similar considerations to the ones I developed in Beer's case also apply to Zednik's interpretation of the A-not-B error model by Thelen and collaborators. Again, I don't see why the contribution here must be strictly understood in terms of functional decomposition of a complex task into low-level (perceptual and motor) activities. Analyzing this case exceeds my purposes here.

5. In line with my remarks, I would agree with rejecting the dichotomy framed in this way, but my proposal is very different from the authors'.

References

Abrahamsen, Adele, and William Bechtel. 2006. "Phenomena and mechanisms: Putting the symbolic, connectionist, and dynamical systems debate in broader perspective." Pp. 159-185 in *Contemporary Debates in Cognitive Science*. Edited by Robert Stainton. Malden, Ma.: Blackwell.

Abrahamsen, Adele, and William Bechtel. 2012. "From reactive to endogenously active dynamical conceptions of the brain." Pp. 329-366 in *Philosophy of Behavioral Biology*. Edited by Kathryn Plaisance and Thomas Reydon. Amsterdam: Springer.

Aminoff, Elissa, Daniela Balslev, Paola Borroni, Ronald Bryan, Elizabeth Chua, Jasmin Cloutier, Emily Cross, Trafton Drew, Chadd Funk, Ricardo Gil-da-Costa, Scott Guerin, Julie Hall, Kerry Jordan, Ayelet Landau, Istvan Molnar-Szakacs, Leila Montaser-Kouhsari, Jonas Olofsson, Susanne Quadflieg, Leah Somerville, Joselyn Sy, Lucina Uddin, and Makiko Yamada. 2009. "The landscape of cognitive neuroscience: Challenges, rewards, and new perspectives." Pp. 1253-1260 in *The Cognitive Neurosciences IV*. Edited Michael Gazzaniga. Cambridge, Ma.: MIT Press.

Bechtel, William. 1998. "Representations and cognitive explanations: Assessing the dynamicist's challenge in cognitive science." *Cognitive Science* 22: 295-318.

Bechtel, William. 2001. "The compatibility of complex systems and reduction." *Minds and Machines* 11: 483-502.

Bechtel, William. 2002. "Aligning multiple research techniques in cognitive neuroscience." *Philosophy of Science* 69: 48-58.

Bechtel, William. 2008. *Mental Mechanisms*. London: Routledge.

Bechtel, William and Robert Richardson. (1993) 2010. *Discovering Complexity*. Cambridge, Ma.: MIT Press.

Beer, Randall. 1996. “Toward the evolution of dynamical neural networks for minimally cognitive behavior.” Pp. 421-429 in *From Animals to Animats*. Vol. 4. Edited by Pattie Maes, Maja Matarić, Jean-Arcady Meyer, Jordan Pollack, and Stewart Wilson. Cambridge, Ma.: MIT Press.

Beer, Randall. 1997. “The dynamics of adaptive behavior: A research program.” *Robotics and Autonomous Systems* 20: 257-289.

Beer, Randall. 2000. “Dynamical approaches in cognitive science.” *Trends in Cognitive Sciences* 4: 91-99.

Beer, Randall. 2003. “The dynamics of active categorical perception in an evolved model agent.” *Adaptive Behavior* 11: 209-243.

Bickle, John. 2006. “Reducing mind to molecular pathways: Explicating the reductionism implicit in current cellular and molecular neuroscience.” *Synthese* 151: 411-434.

Bickle, John, and Valerie Hardcastle. 2012. “Philosophy of neuroscience.” *Elsevier Life Sciences Reviews*. doi: 10.1002/9780470015902.a0024144.

Buzsáki, Gyorgy. 2006. *Rhythms of the Brain*. Oxford: Oxford University Press.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. Cambridge, Ma.: MIT Press.

Chemero, Anthony, and Michael Silberstein. 2008. “After the philosophy of mind.” *Philosophy of Science* 75: 1-27.

Cooper, Richard, and Tim Shallice. 2010. “Cognitive neuroscience: The troubled marriage of cognitive science and neuroscience.” *Topics in Cognitive Science* 2: 398-406.

Craver, Carl. 2005. “Beyond reduction: Mechanisms, multifield integration, and the unity of science.” *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 373-396.

Craver, Carl. 2006. “What mechanistic models explain.” *Synthese* 153: 355-376.

Craver, Carl. 2008. “Physical law and mechanistic explanation in the Hodgkin and Huxley model of the action potential.” *Philosophy of Science* 75: 1022-1033.

Cummins, Robert. 1975. “Functional analysis.” *Journal of Philosophy* 72: 741-764.

Dale, Rick. 2008. “The possibility of a pluralist cognitive science.” *Journal of Experimental and Theoretical Artificial Intelligence* 20: 155-179.

Dale, Rick, Eric Dietrich, and Anthony Chemero. 2009. “Explanatory pluralism in cognitive science.” *Cognitive Science* 33: 739-742.

Deco, Gustavo, Giulio Tononi, Melanie Boly, and Morten Kringelbach. 2015. “Rethinking segregation and integration: Contributions of whole-brain modelling.” *Nature Reviews Neuroscience* 16: 430-439.

Dietrich, Eric, and Arthur Markman. 2001. “Dynamical description versus dynamical modeling.” *Trends in Cognitive Sciences* 5: 332.

Dupré, John. 2013. “Living causes.” *Aristotelian Society Supplementary* 87: 19-37.

Engel, Andreas, Pascal Fries, and Wolf Singer. 2001. “Dynamic predictions: Oscillations and synchrony in top-down processing.” *Nature Reviews Neuroscience* 2: 704-716.

Engel, Andreas, Karl Friston, Scott Kelso, Peter König, Ilona Kovács, Angus MacDonald III, Earl Miller, William Phillips, Steven Silverstein, Catherine Tallon-Baudry, Jochen Triesch, and Peter Uhlhaas. 2010. “Coordination in behavior and cognition.” Pp. 267-299 in *Dynamic Coordination in the Brain: From Neurons to Mind*. Edited by Christoph von der Malsburg, William Phillips, and Wolf Singer. Cambridge, Ma.: MIT Press.

Freeman, Walter. 2005. "A field-theoretic approach to understanding scale-free neocortical dynamics." *Biological Cybernetics* 92: 350-359.

Gazzaniga, Michael, Karl Doron, and Chadd Funk. 2009. "Looking toward the future: Perspectives on examining the architecture and function of the human brain as a complex system." Pp. 267-299 in *The Cognitive Neurosciences IV*. Edited by Michael Gazzaniga. Cambridge, Ma.: MIT Press.

Gervais, Raoul. 2015. "Mechanistic and non-mechanistic varieties of dynamical models in cognitive science: Explanatory power, understanding, and the 'mere description' worry." *Synthese* 192: 43-66.

Giunti, Marco. 1997. *Computation, Dynamics, and Cognition*. New York: Oxford University Press.

Glennan, Stuart. 2002. "Rethinking mechanistic explanation." *Philosophy of Science* 69: 342-353.

Glennan, Stuart. 2005. "Modeling mechanisms." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 443-464.

Harnad, Stevan. 1987. "Introduction: Psychophysical and cognitive aspects of categorical perception." Pp. 1-25 in *Categorical Perception: The Groundwork of Cognition*. Edited by Stevan Harnad. Cambridge: Cambridge University Press.

Huster, René, Stefan Debener, Tom Eichele, and Christoph Herrmann. 2012. "Methods for simultaneous EEG-fMRI: An introductory review." *The Journal of Neuroscience* 32: 6053-6060.

Izhikevich, Eugene. 2007. *Dynamical Systems in Neuroscience*. Cambridge, Ma.: MIT Press.

Kaplan, David, and William Bechtel. 2011. “Dynamical models: An alternative or complement to mechanistic explanations?” *Topics in Cognitive Science* 3: 438-444.

Kaplan, David, and Carl Craver. 2011. “The explanatory force of dynamical and mathematical models in neuroscience.” *Philosophy of Science* 78: 601-627.

Kelso, Scott. 1995. *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, Ma.: MIT Press

Koertge, Noretta. 1992. “Explanation and its problems.” *British Journal for the Philosophy of Science* 43: 85-98.

Matthewson, John. 2011. “Trade-offs in model-building: A more target-oriented approach.” *Studies in History and Philosophy of Science* 42: 324-333.

Mitchell, Melanie. 2009. *Complexity: A Guided Tour*. Oxford: Oxford University Press.

Rosa, Maria, Jean Daunizeau, and Karl Friston. 2010. “EEG / fMRI integration: A critical review of biophysical modeling and data analysis approaches.” *Journal of Integrative Neuroscience* 9: 453-476.

Revonsuo, Antti. 2001. “On the nature of explanation in the neurosciences.” Pp. 45-69 in *Theory and Method in the Neurosciences*. Edited by Peter Machamer, Peter McLaughlin, and Rick Grush. Pittsburgh: University of Pittsburgh Press.

Rodriguez, Eugenio, Nathalie George, Jean-Philippe Lachaux, Jaques Martinerie, Bernard Renault, and Francisco Varela. 1999. “Perception's shadow: Long-distance synchronization of human brain activity.” *Nature* 397: 430-433.

Rosenbaum, David. 1998. “Is dynamical systems modeling just curve fitting?” *Motor Control* 2: 101-104.

Ross, Lauren. 2015. “Dynamical models and explanation in neuroscience.” *Philosophy of Science* 82: 32-54.

Schöner, Gregor, and Hendrik Reimann. 2009. “Understanding embodied cognition through dynamical systems thinking.” Pp. 450-473 in *The Routledge Companion to Philosophy of Psychology*. Edited by John Symons and Paco Calvo. New York: Routledge.

Silberstein, Michael, and Anthony Chemero. 2013. “Constraints on localization and decomposition as explanatory strategies in the biological sciences.” *Philosophy of Science* 80: 958-970.

Slocum, Andrew, Douglas Downey, and Randall Beer. 2000. “Further experiments in the evolution of minimally cognitive behavior.” Pp. 430-439 in *From Animals to Animats*. Vol. 6. Edited by Jean-Arcady Meyer, Alain Berthoz, Dario Floreano, Herbert Roitblat, and Stewart Wilson. Cambridge, Ma.: MIT Press.

Smith, Linda and Esther Thelen. 2003. “Development as a dynamic system.” *Trends in Cognitive Sciences* 7: 343-348.

Stepp, Nigel, Anthony Chemero, and Michael Turvey. 2011. “Philosophy for the rest of cognitive science.” *Topics in Cognitive Science* 3: 425-437.

Stewart, Lauren, and Vincent Walsh. 2006. “Transcranial magnetic stimulation in human cognition.” Pp. 1-26 in *Methods in Mind*. Edited by Carl Senior, Tamara Russell, and Michael Gazzaniga. Cambridge, Ma.: MIT Press.

Strogatz, Steven. 1994. *Nonlinear Dynamics and Chaos*. Reading: Addison-Wesley.

Sullivan, Jaqueline. 2009. “The multiplicity of experimental protocols.” *Synthese* 167: 511-539.

Thelen, Esther, Gregor Schöner, Christian Scheier, and Linda Smith. 2001. “The dynamics of embodiment.” *Behavioral and Brain Sciences* 24: 1-86.

van Gelder, Tim. 1997. “Dynamics and cognition.” Pp. 421-450 in *Mind Design II: Philosophy, Psychology, Artificial Intelligence*. Edited by John Haugeland. Cambridge, Ma.: MIT Press.

van Gelder, Tim. 1998. “The dynamical hypothesis in cognitive science.” *Behavioral and Brain Sciences* 21: 615-665.

van Gelder, Tim, and Robert Port. 1995. “It's about time.” Pp. 1-43 in *Mind as Motion*. Edited by Robert Port and Tim van Gelder. Cambridge, Ma.: MIT Press.

van Leeuwen, Marco. 2005. “Questions for the dynamicist.” *Minds and Machines* 15: 271-333.

Walmsley, Joel. 2008. “Explanation in dynamical cognitive science.” *Minds and Machines* 18: 331-348.

Weiskopf, Daniel. 2011. “Models and mechanisms in psychological explanation.” *Synthese* 183: 313-338.

Woodward, James. 2015. “Explanation in neurobiology: An interventionist perspective.” Forthcoming in *Integrating Psychology and Neuroscience: Prospects and Problems*. Edited by David Kaplan. Oxford: Oxford University Press.

Zednik, Carlos. 2011. “The nature of dynamical explanation.” *Philosophy of Science* 78: 236-263.