How Not to Argue about the Compatibility of Predictive Processing and 4E Cognition

Yavuz Recep Başoğlu*

Received: 16 February 2020 / Revised: 26 May 2020 / Accepted: 9 June 2020

Abstract: In theories of cognition, 4E approaches to cognition are seen to refrain from employing robust representations in contrast to Predictive Process, where such posits are utilized extensively. Despite this notable dissimilarity with regard to posits they employ in explaining certain cognitive phenomena, it has been repeatedly argued that they are in fact compatible. As one may expect, these arguments mostly end up contending either that Predictive Process is actually nonrepresentational or that 4E approaches are representational. In this paper, I will argue that such arguments are inadequate for the indicated purpose for several reasons: the variety of representational posits in Predictive Process, the diverse attitudes of practitioners of 4E approaches toward representations and the unconstrained use of the term “representation” in cognitive science. Hence, here I will try to demonstrate that any single argument, if it depends on representational 4E approaches or nonrepresentational Predictive Process, falls short of encompassing this heterogeneity in pertinent debates. Then, I will analyze similar arguments provided by Jacob Hohwy and Michael Kirchhoff to illustrate how destructive this seemingly ordinary criticism is.
Keywords: 4E cognition, embodied cognition, free-energy principle, mental representation, predictive processing, representation wars.

1. Introduction

While 4E approaches to cognition (i.e., embodied, extended, embedded, and enactive; one might also include ‘situated,’ ‘distributed,’ ‘affective,’ ‘interactive,’ ‘extensive,’ etc.) (4E cognition henceforth), which are also gathered under the generic name “second-generation cognitive science” (Lakoff and Johnson 1999, 78), have enjoyed their heyday over the last decades, another idea has also recently started to excite the cognitive scientist: Predictive Processing (PP henceforth) as a particularly promising neuro-computational framework purporting to account for a uniform understanding of the mechanisms underlying cognition.

Whereas practitioners of 4E cognition are rigorously at odds with traditional cognitivists who recruit internal symbolic representations ubiquitously in their explanations of cognitive phenomena, PP is thought to make use of alleged representations nontrivially to use an internal model of the external world. For this reason, prima facie, these two theories of cognition seem to be incompatible. That is, at least one must be false. Nevertheless, philosophers of mind and cognitive scientists have been at pains to prove that PP and 4E cognition are compatible (e.g., Clark 2013; 2015a; Kirchhoff 2018; Gladziejewski 2017; Hohwy 2018). Hence, to establish the compatibility of PP and 4E cognition in terms of representations, scholars mostly argue either that 4E is representational or that PP is nonrepresentational to leave no fundamental difference that might suggest that these theories of cognition are, in fact, incompatible.

In section 2 of this paper, I first provide a brief and selective overview of PP and 4E cognition and expound how the question of their compatibility boils down to questions of the ontological nature of representational posits deployed by these theories. In section 3, I try to explicate the general schema of arguments advanced when arguing for the compatibility of PP and 4E cognition in terms of representations. I shall argue that any argument contending that PP and 4E cognition are compatible is doomed to fail when it depends on either representational 4E cognition or nonrepresentational PP.

Organon F 28 (4) 2021: 777–801
This is the case because the term “representation” as employed by the cognitive scientist, the attitudes of 4E cognition proponents toward representations, and the kinds of representational posits applied in the PP framework are not homogenous. It is thus not possible for a single argument concerning the existence of representations used in any of these theories to embrace this heterogeneity. It is no longer obscure that the unconstrained use of the term representation creates a flexibility to read any theory in representational or nonrepresentational terms. This trivializes the representationalism vs. anti-representationalism debate and renders them “for the sake of appearance” (Haselager et al. 2003, 21). This, I shall argue, turns out to be more daunting in debates on the relation between 4E cognition and PP. In section 4, I analyze the arguments of both Hohwy (2018) and Kirchhoff (2018) for the compatibility of PP and 4E cognition and show that both arguments employ the flawed argument schema scrutinized in section 3. Thereafter, the main task of this section is to illustrate how Hohwy and Kirchhoff’s arguments fail to be conclusive in establishing the compatibility of PP and 4E cognition. This illustration aims to emphasize how significant the seemingly simple criticism voiced in section 3 is and how destructive it can be when ignored.

2. 4E cognition, predictive processing and the representation wars

Ushering in a whole new era in the philosophy of mind and cognitive science, 4E cognition arose as a result of growing discontent with traditional cognitivism’s claim that cognition is brain-bound or intracranial and is a kind of computation realized by syntactically manipulating symbolic mental representations (see Fodor 1975 and Pylyshyn 1984). Lakoff and Johnson’s (1980) idea of the metaphoric use of language as not peripheral to cognition but as fundamental to how one conceptualizes the world, including time, space, feelings, etc., and Varela et al.’s (1991) emphasis on embodied action especially initiated this approach. New studies and developments in robotics (e.g., Brooks 1991) and dynamical systems (e.g., Thelen and Smith 1994) have also accelerated the advent of 4E cognition. While it seems quite futile
to attempt to provide a uniform definition encompassing the whole research program for 4E cognition as is made evident in section 3, its most distinctive characteristics can be given as follows: it prioritizes the idea of the body and environment as constitutive in cognition rather than as just peripheral to certain cognitive processes. This means that in favor of an embodied and extended approach, 4E cognition undermines the intracranialism (e.g., Adams and Aizawa 2008) associated with traditional cognitivism, which asserts that cognition, as a matter of empirical fact, is brain-bound or intracranial. 4E cognition also accentuates nonlinear constant dynamical coupling between action and perception, and among the brain, body and environment, which purports that intelligent behavior is not driven by a certain form of computation but results from this dynamical coupling. Thus, the body and the environment not only causally contribute to cognition but also play a constitutive role, and the brain, the body, and the environment are not separated from each other, but constitute one continuously interacting system.

Since the body and environment have constitutive roles in cognition and are in constant dynamical coupling, they are in a position to provide the information required for cognition on the fly. 4E cognition thus defends the notion that representations that are supposed, according to the cognitivist, to carry such information are mostly not needed nor desirable in explaining cognitive phenomena. Thereupon, most importantly for the issue at hand, 4E cognition researchers (e.g., Goldman 2012; Van Gelder 1995; Noë 2004) aim to minimize if not altogether to dispense with the elementary role of amodal, action-neutral, quasi-linguistic, contentful, and symbolic inner representations of cognitivist understandings of the mind. This tendency among practitioners of 4E cognition quickly escalated ‘the representation wars’ between representationalists and anti-representationalists, and even after almost three decades of discussion, these representation wars have not yet been settled.

In the thick of the war, PP posits that to engage the world around it, the brain, which uses Bayesian reasoning, continually predicts incoming

---

1 The term “body” refers to extracranial parts of the body or to the body minus the brain in this context.

2 The words ‘mind,’ ‘brain,’ and ‘agent’ in this discussion appear to be synonymous. Their use is a matter of taste. Slightly different connotations relative to their
sensory data by means of its top-down hierarchical architecture, which models hidden nested causal complexities of the external world. The predictions generated by these top-down processes are compared to incoming sensory data, and error messages are generated when there is a mismatch between top-down predictions and the incoming sensory data. Generated error messages propagate up the hierarchy to adjust the internal model of the environment, which in turn helps minimize errors in predictions made at lower levels of the hierarchy and hence generates more accurate predictions for following iterations. To minimize prediction errors, the internal model of nested causal complexities of the environment can be updated and adjusted as just explained. The brain’s other strategy in performing prediction error minimization (PEM henceforth) involves moving its sensory organs in a certain way to fit incoming sensory data to the prediction already generated by the system’s hierarchical architecture. In so doing, the brain proactively samples sensory data. This particular mode of PEM is called active inference, which emphasizes continuous dynamical coupling between perception and action and between the environment and the agent. Thus, “perceiving and acting are but two different ways of doing the same thing” (Hohwy 2013, 76).

To appreciate the peculiarities of PEM, one must also place it in the context of the Free Energy Principle (FEP henceforth), according to which “any self-organizing system that is at equilibrium with its environment must minimize its free energy” (Friston 2010, 127). This means that an organism strives to maintain homeostasis in the face of continuously changing environmental factors. The organism avoids situations that are unexpected or surprising (or more technically showing higher levels of surprisal) for its phenotype. In rendering the PEM strategy of PP a special version of the FEP, top-down predictions are generated in accordance with bodily and environmental states, which are expected not to exhibit higher surprisal relative to its phenotype (e.g., being out of water has high surprisal relative to the fish phenotype but not to the human phenotype, and accordingly, due to the difference between their phenotypes, the fish expects to be in water, whereas the human does not.) This suggests that “all aspects of perception and

original fields are ignorable for the issue at hand. Throughout the paper, I respect the choices of the authors in question and use them accordingly.
cognition then have a foundation in bodily states and movement and purposeful behavior have a foundation in the environment” (Hohwy 2018, 137).

The PEM strategy is then what underlies cognition. In pursuing PEM, the brain can detect the hidden causes that best explain sensory data impinging upon its sensory organs. Thus, the brain is nothing but a continuously working prediction machine (see also Clark 2013, 2016).

Due to the scope of this paper, several aspects of the theories in question are intentionally ignored. However, for the present study, this selective overview of PP and 4E cognition shall demonstrate that proponents of PP have made certain cases to claim that it is compatible with 4E cognition on the grounds that both emphasize dynamical coupling among the brain, body, and environment and give fundamental roles to the body and environment in cognition. “There remains, however, at least one famously vexed issue” (Clark 2016, 291): the issue of representation. Nevertheless, due to its aforementioned 4E-friendly tenets, Clark (2015a) (see also Madary 2015) was expecting PP to bring about peace and to end the representation wars. However, it instead opened a new front.

Alleged representations of the PP framework have been, and still are, central to various discussions, and especially to those regarding the relation between PP and 4E cognition. Two different lines of discussion are still going on: whether representations are involved in cognition and whether representational posits of PP are robust representations. The jury is still out on both discussions. For the latter, among the camps are conservative and radical versions of PP. While supporters of conservative PP argue that what PP construes are robust representations that are employed in a non-trivial way (see Gładziejewski 2015; Hohwy 2014), radical PP argues for a somehow deflated version of representations of PP. Radical PP thus claims either that representational posits in PP are not full-blooded or robust representations like those embedded in cognitivism (Kirchhoff and Robertson 2018) or that at least not all representational posits are robust representations of cognitivist understanding (Clark 2015b; Orlandi 2014). Generally, discussion on the relation between PP and 4E cognition has considered the above division: conservative and radical PP. Interestingly, however, a great many in both camps argue that their version of PP is 4E cognition friendly (e.g., Gładziejewski 2017; Clark 2016; Kirchhoff 2018; Hohwy 2018).
3. Representational 4E cognition versus nonrepresentational PP

In arguing for or against the compatibility of PP and 4E cognition in terms of representational posits, the typical strategy adopted is as follows. First, one demonstrates the tenets achieved by the PP framework, which match the characteristics of 4E cognition such as those briefly overviewed in section 2. However, this forces her to make one last decision between a representational reading or nonrepresentational reading of both theories, since while one theory uses explicitly representational terminology, the other abstains from such terminology as much as it can. Thus, second, she focuses on certain properties (e.g., decoupleability) of the representational posits, which are called on to perform 4E cognition-friendly tenets and to determine whether such properties qualify them as robust representations depending on a specific definition such as Ramsey’s (2007). Through this line of reasoning, one discusses the existence and employment of representations. Thus, the following is the argument schema of this way of arguing:

(Non) Representationality Argument:
1) With its means and strategy, PP is able to realize 4E cognition-relevant cognitive tenets.
2) PP is not representational (or 4E cognition is representational) because representational posits needed to embrace these tenets of 4E cognition are (or are not) robust representations according to such a definition.
3) Therefore, PP and 4E cognition are compatible.

I shall argue that any (non) representationality argument is inefficient because we do not have a uniform idea of representationalism or anti-representationalism to argue that the second premise is true. The need for such uniformity and the problems caused by a lack of it have already been repeatedly pointed out in this debate (Haselager et al. 2003; Svensson and Ziemke 2005; Ramsey 2007). In what follows, it will be shown that this problem is more virulent than once thought for the issue at hand and in fact renders all (non) representationality arguments for the compatibility of PP and 4E cognition invalid.
The first reason relates to the fact that the ontological nature and representational status of posits of the PP framework are still highly controversial even among the most prominent proponents of PP. As to the status of representational posits of the PP framework, most involved in this discussion, whether conservative or radical, accept Ramsey’s “job description challenge” as the criterion to be a robust representation. This challenge aims to constrain the liberal use of this term and to prevent the term from being trivialized. (Ramsey 2007) asks philosophers if the internal states they wish to call “representations” satisfy this challenge. If they do not, Ramsey suggests that they are not representations proper. Although job description challenge is a theoretical term, it depends on the daily use of the term “representation”: to describe the representational function of representational posits in a cognitive system (Ramsey 2007). Thus, theories that posit representations are under an obligation to genuinely demonstrate the sense in which these internal states function as representations or as stand-ins for the external state of affairs. In his own words, he asks for “a job description that tells us what it is for something to function as a representation in a physical system” (Ramsey 2007, 27). He argues that posits of new generation theories in the sciences of mind (such as Dretske 1988) fail to meet this challenge and do not quality as representations, but he does not comment on PP’s representational posits.

As the first aspect of the problem, proponents of PP who accept Ramsey’s definition as the criterion to qualify as a representation conclude different results. On one hand, some (e.g., Gladziejewski 2015; Kiefer and Hohwy 2017) argue that according to Ramsey’s definition, the representational posits of PP satisfy this challenge and serve as structural representations. Gladziejewski (2015), for example, employs a compare-to-prototype strategy also offered by Ramsey as a way to describe the role of internal states. Through this strategy, one finds a structure in daily life referred to as representation in an uncontroversial way and finds that this structure functions in a similar way to the internal states that one wishes to call representation. Gladziejewski argues that internal states in PP resemble cartographic maps in their job descriptions. Thus, he concludes that such

---

3 I would like to thank the anonymous referee for pressing me on this point.
posits are “(1) structural representations that (2) guide the actions of their users, (3) do so in a detachable way and (4) allow their users to detect representational errors” (Gładziejewski 2015, 567). On the other hand, some believe that representational posits in PP cannot be called representations (e.g., Kirchhoff and Robertson 2018, Orlandi 2014). Proponents of radical PP draw attention to generative models of the lower-level domains (e.g., perception) and claim that such architectures are, as it were, model-free and incapable of, say, detaching from what they stand for. One reason for this disagreement relates to the fact that radical and conservative PP focuses on different parts of the system, which will be clearer in the next section. As the second aspect, for this reason, some (e.g., Dolega 2017) argue that Ramsey’s challenge is inadequate for PP that makes use various generative models with a multilayered architecture. With this challenge, Ramsey seems to tie representations to their functions. In his treatment, among the many other properties, the function of internal states emerges as the decisive property. Such complex structures and components of PP allow for different functional readings as seen. Consequently, Dolega rejects Ramsey’s criterion because it is not a suitable challenge for PP’s complex and multilayered architecture, arguing that the status of such posits should be discussed while considering the content rather than representational functions.

Thus, this discussion of job description challenge reveals that, let alone discussions of whether PP is representational, there is no consensus on some prior points: (1) which properties of representational posits should be central to the evaluation of generative models of PP, (2) whether the job description challenge is suitable for PP’s generative model, and (3) which parts of the system should be considered when evaluating PP. While nothing is for sure, the fact is that these representational posits of the PP framework are not the same as those of the cognitivist tradition. There are still broad disagreements about which definition is to be applied to the representational posits of PP and whether they are robust representations according to any given definition.

In addition to ambiguities as to how to evaluate the generative model, there is more than one type of posit of PP that might be considered a robust representation. As Gładziejewski (2017) rightly appreciates, in PP there are at least four more types of posits that might also earn the status of robust
representation: the generative model, sensory signals, prediction error signals, and precision estimators. Orlandi (2014) argues that other posits such as prediction signals, and error signals do not satisfy Ramsey’s job description challenge. In debates on representational posits of PP, theorists generally only mention the generative model while ignoring the other posits, which also deserve closer scrutiny and might be decisive in this debate.

The most significant issue anyone discussing the compatibility of 4E cognition and PP in terms of representations must keep in mind is that the term ‘representation as rejected by the proponents of 4E cognition’ is not a homogenous term. It comes with two aspects. As the first aspect, 4E cognition theorists argue against representations on the basis of different definitions and take distinct attitudes toward them. As the second aspect, proponents of 4E cognition raise various challenges against representations. For instance, Chemero (2009) argues against representations on the basis of Millikan-style representations (Millikan 1984; 1993), which he thinks carry the general characteristics of traditionally employed robust representations. He argues that action-oriented (Clark 1997), pushmi-pullyu (Millikan 1995), indexical-functional (Agre and Chapman 1987) and emulator representations (Grush 1997; 2004) are robust representations. In contrast, Gallagher (2008; 2017) renounces the representational status of these representations on the basis of Rowlands’ definition (Rowlands 2006), which interestingly echoes Millikan-style representations. Surprisingly, Thompson, coauthor of The Embodied Mind (Varela et al. 1991), which is regarded the urtext of 4E cognition, expressed his sympathy for the emulator account of mental imaginary in (Foglia and Grush 2011), which makes use of emulator representations (Grush 1997; 2004). He wrote that “[I] see it as a friendly supplement to my remarks about sensorimotor processes in mental imagery (Thompson 2011, 194)” while granting its status as a form of representation. Thus, these three thoroughgoing proponents of 4E cognition have completely disparate attitudes toward a certain sort of representational posit.

For the second aspect, Varela et al. (1991) argue against the traditional cognitive function of the mind. They suggest that the function of cognition

---

4 I by no means aim to be exhaustive in revealing the diversity of versions of 4E cognition. It suffices for the present purposes to show how many different versions there could be.
is to engage the world and not to think or build an action-neutral duplicate of the world. Accordingly, they are against the traditional function of representations of creating an action-neutral internal model of the external world. Barsalou (1999; 2008) points out what is called “the symbol grounding problem” and rejects the amodality of representations of traditional cognitivist understanding. Goldman (2012) proposes a somehow weakened version of embodied cognition and accepts what he has dubbed “bodily formatted representation,” whose content is information about bodily states collected by means of interoception, and for him, this is how the body earns its constitutive role in cognition. Hutto and Myin (2013; 2018) raise the content determination problem (or what they call “the hard problem of content”), which refers to the impossibility of physical states (in this case, representations) happening to bear any content whatsoever. Thus, what they reject is to employ any content-bearing representational posit.

Hence, any work not speaking to the concerns mentioned above is doomed to be inconclusive, as neither “representation” nor anti-representationalism is homogenous. Thence, to what extent representational posits of PP satisfy the various criticisms raised by proponents of 4E cognition should be central to discussions of the compatibility of PP and 4E cognition. Any (non) representationality argument falls short of revealing this relation. Whenever one accepts a definition to argue for the (non) existence of representation in PP or 4E cognition, there will always be some equally consistent counter definitions that might be employed against him. Whenever one argues for or against the compatibility of PP and 4E cognition on the basis of the (non) existence of representations in either PP or 4E cognition, one will be arguing against or for only a specific version of 4E cognition or anti-representationalism.

It has been already argued that “[w]ithout a properly constrained notion of representation, the debate between representationalists and anti-representationalists is bound to remain a debate for the sake of appearance” (Haselager et al. 2003, 21) because the liberal and unconstrained use of representation as a term in both camps makes any debate between them futile. In this section, I hope to have shown that in the case of representation wars between PP and 4E cognition, the problem is deeper and more fundamental than in any representationalism vs. anti-representationalism debate.
In what follows, I aim to illustrate how destructive this criticism is by analyzing arguments for the compatibility of PP and 4E cognition provided in (Hohwy 2018) and (Kirchhoff 2018).

4. An analysis of Hohwy and Kirchhoff’s arguments

*The Oxford Handbook of 4E Cognition* (Newen et al. eds. 2018) contains two seminal articles arguing that PP and 4E cognition are indeed compatible. In chapter 7 of this book, Hohwy (2018) spells out how PP, about which he has been theorizing (e.g., Hohwy 2013; 2014; 2017), is congenial to 4E cognition “only because 4E cognition, rightly understood, is nothing but representation and inference” (Hohwy 2018, 130). In chapter 12, Kirchhoff (2018) contends that PP is compatible with the nonrepresentational thesis of 4E cognition. Thus, both argue that PP and 4E cognition are congenial, but for different reasons. One suggests that 4E cognition is representational while the other proposes that PP is not representational. I shall argue here that both fail to be conclusive. In (Hohwy 2018) and (Kirchhoff 2018), the term “representation” is used ambiguously, and they do not explain what they mean by this vague term. Each focuses on a specific part of the hierarchical architecture of PP and on a particular set of properties of such posits and decisively arrives at a final verdict: that PP is nonrepresentational or that 4E cognition is representational. That is, both present instances of (non) representationality arguments explicated in section 3.

Let us first contemplate the argument made in (Hohwy 2018). The argument schema presented is as follows:

1) PP is necessarily representational.

2) 4E cognition, correctly understood, is representational because PP, given its representational nature, can “encompass phenomena highlighted in debates on 4E cognition” (Hohwy 2018, 130).

Kirchhoff (2018) additionally argues that the PP framework is compatible with the constitutive, cognitive-affective inseparability, and metaplasticity theses of what he thinks 4E cognition is, and Hohwy (2018) also argues for the inferentiality of 4E cognition. In this section, considering the scope of the present study, I shall only focus on their representationality arguments.
3) Therefore, 4E cognition and PP are compatible.

To argue for the first premise, he demonstrates how representations are, as it were, *sine qua non* for the brain’s PEM strategy. He does not explicitly specify any cognitive domain or any parts of the hierarchical architecture and appears to deal with the whole multilayered structure. Representational posits of PP play the functional role of mirroring nested causal complexities of the external world. The better the external world is represented, the more accurate predictions generated by the model will be, making such representations essential for PP. For the second premise, he points out that due to the FEP (see Friston 2010; Friston and Stephan 2007), the surprise is relative to the model formed directly by the organism’s phenotype. This endows bodily states and the environment with central roles in cognition. The content is then organism-salient, and the body and natural environment of the organism are not trivial but play fundamental roles. The content of those posits speaks to the tenets of 4E cognition. Thus, since he takes representations for granted from the first premise, he declares that 4E cognition is representational depending on the FEP, which renders PP compatible with essential tenets of 4E cognition. He takes the representational function of the generative model and draws a conclusion about representational content congenial to 4E cognition.

The first concern as to the validity of his argument stems from what Hohwy means by the term “representation,” which he argues is necessary for PP. Hohwy mentions two hallmarks of representations: action-guidance and detachability (Hohwy 2018, 135). He derives the functional property of action-guidance from active inference, but he does not elaborate more on in what sense those representational posits are ‘detached’ or even on what it means for a posit to be ‘detached.’ It is highly controversial whether these two properties turn any posits into representations. When arguing that PP is representational, Gładziejewski (2015, 12) claims that ‘acting-guiding’

---

6 This point is more crucial than it seems because in the literature one can find various understandings of “detachment” and “decoupleability,” which might be decisive in understanding what kinds of representations Hohwy refers to. For the present study, it is enough to raise this question and not to go into detail to avoid further discussion. See Rowlands (2012) and Gallagher (2017, 88-103) for more discussion on the significance of these terms in the representation wars.
and ‘detachability’ are not jointly sufficient to be counted as representation. He seems to refer to a pretheoretic understanding of internal representations (i.e., mirroring nature) and does not go into theoretical details to prove that these are robust representations according to any given definition.

Next, even if one assumes that his premises are correct, one must raise the question of which of the 4E notions Hohwy’s framework is compatible with. Hohwy acknowledges that he refers to “somehow deflated 4E notions” (Hohwy 2018, 130). However, details of these “somehow deflated 4E notions” remained obscure. In a paper of this scope, it is not feasible to attempt to review all versions of 4E cognition and to find the ones Hohwy refers to. Nevertheless, certain examples can be analyzed to illustrate the possible positions Hohwy might not have taken. Then, first, it is clear that for Hohwy these posits are content-bearing states because he already argued that the content of a posit is organism-salient. Thus, his proposal does not carry any kinship to Hutto and Myin (2013) because for them, a great many number of cognitive phenomena are contentless. His account also stands in strong opposition to those for whom just a covariance relation must suffice for at least some cognitive tasks (especially lower-level ones) such as Gallagher (2017), Chemero, (2009), Noë (2004), and Van Gelder (1995). Though action-guiding, Hohwy’s posits have the function of building a rich and fragile representation of the external world, which, without further ado, situates him against a variety of enactivists such as Varela et al. (1991).

It is somehow painless to point out whose representational posits are not compatible with Hohwy’s. However, if his argument is not that PP is compatible with what I call 4E cognition, what he refers to by “somehow deflated 4E cognition” must be found in the literature. This seems, however, slightly more challenging. He names in the article in question some 4E cognition proponents such as Varela (1991), Gallagher (2005), and Thompson (2007) with whom his framework might “have contact” and others such as

---

Note that while the main tendency among contenders of debates on representations is to think that a covariance relation is not enough to provide natural content for representational posits, some can still argue that it is. Thus, to avoid further complications and discussions resulting from these different understandings of philosophers, I refer to those who use the term “contentless” and to those who prefer to use the term “just covariance” separately.
Van Gelder (1995) to whom his theory “speaks.” Above I have already shown that his representational posits are incompatible with those in these accounts. Possible candidates would be among those 4E notions deploying representational terminology explicitly such as Clark (2008), Wheeler (2005), Goldman (2012), Goldman and de Vignemont (2009), Bickhard (2000), Barsalou (1999), Mandik (2005), Prinz (2009), and Hutchins (1995). However, one must be careful in claiming that (Hohwy 2018)’s account is compatible with these accounts based on the fact that both use explicit representational terminology because, for example, Wheeler (2005; 2008) does away with decoupleability as a criterion for minimal representation. If Hohwy’s ‘detachment’ means ‘decoupleability’ in the sense that a representational posit can also perform its function of guiding intelligent behavior (or action-guiding as stated in (Hohwy 2018)) even when the represented extracranial target is absent, then this jeopardizes their compatibility at least for some lower level cognitive phenomena. That is, Wheeler can explain certain cognitive phenomena with posits that are not decoupleable with their target, but Hohwy (2018) cannot. Furthermore, since it is clear that Hohwy’s posits carry extracranial bodily content, this might be compatible with Mandik (2005), but whether these posits have the bodily format (not only content)9 required by Goldman (2012), Goldman and de Vignemont (2009), Barsalou (1999), etc. is not palpable in (Hohwy 2018).10

---

8 He also mentions (Clark 1997) and (Aydede and Robbins eds. 2009). However, the central arguments in these works concern intracranialism, not representationalism. Since the present study is exclusive to the representationality discussion, both works are intentionally excluded.

9 For a detailed discussion of and substantial criticism concerning “bodily content” and “bodily format,” see Hutto (2013) and Goldman and de Vignemont (2009), and for a more general discussion of bodily representations and their controversial status in the 4E tradition, see Alsmith and de Vignemont (2012).

10 However, while in (Hohwy 2018) it is hardly possible to understand whether these representational posits are of bodily formats, other texts on PP in the literature reveal that such posits can have not only bodily content but also bodily formats. For example, Gladziejewski wrote that “generative models [...] could bring about patterns of neural activity that resemble those that accompany perception and action” (2017, 106) when performing certain detached cognitive tasks such as imagining, counterfactual reasoning, dreaming, etc.
We are thus only left with a limited number of weakened versions of 4E cognition. If he bites the bullet and reduces his claim to cover only some highly deflated versions of 4E cognition or maybe even only the version of 4E cognition described in (Hohwy 2018), then he faces the risk of trivializing his position in this debate, because what is the point of claiming that versions of 4E cognition that accept representations are representational? He might also be trivializing 4E cognition in general, because I see no reason why an embodied cognitive scientist would accept those overtly representational accounts *qua* accounts of 4E cognition.\(^{11}\)

Please note, however, that the point of the argument is to demonstrate how insufficient arguments for the compatibility of (conservative) PP and 4E cognition are if they depend on representational 4E cognition and not that Hohwy’s framework is only compatible with certain deflated notions of 4E cognition. This is already rightly stated by Schlicht in his critical note to (Hohwy 2018): “PEM can only be compatible with the moderate 4E approaches” (Schlicht 2018, 219).

Let us now turn to Kirchhoff’s argument that PP is compatible with 4E cognition because PP can embrace the nonrepresentational thesis of 4E cognition. The nonrepresentational thesis claims that “[t]he sensorimotor profile of organisms is sufficient for at least some kinds of cognitive activities, thus replacing the need for organisms to construct complex internal mental representations of the outside environment” (Kirchhoff 2018, 244). As the saying goes, one man’s modus ponens is another’s modus tollens. Kirchhoff’s argument touches on the same point as Hohwy’s but from a different perspective. His first premise points out active inference and the FEM context, which emphasizes the continuous coupling of the body, environment, and agent in the PP framework. This idea is a familiar one since this dynamical coupling represents one of the initial ideas of 4E cognition and is supported by almost any defender of 4E cognition (see, e.g., Chemero 2009; Gallagher

---

\(^{11}\) Some of the theories’ positions in the 4E cognition tradition are highly controversial because they explicitly make use of robust representations in their explanations. For instance, Walter renders many of such accounts such as Goldman (2012) as “not embodied at all” (Walter 2014, 246) while Alsmith and de Vignemont consider them “weakly embodied” (2012, 3).
2017). By virtue of active inference, for Kirchhoff, the body and environment become not only coupled but also one interacting system. Thus, the cognitive function of the mind endowed with PP is not to think, but rather to act and engage with the world. With this initial support from active inference, he derives another form of support from surprises relative to a model, which leads him to the same conclusion as Friston that “an agent does not have a model of its world— it is a model” (Friston 2013, 213, emphasis in original). This reading of a model is more inclusive, loose, and pretheoretic; it does not imply strict separation. These peculiarities of PP are compatible with 4E cognition.

His second premise is directed against the argument that to perform PEM, which is also used to minimize nested causal complexity, there must be representations due to self-evidencing.12 Hohwy (2014; 2017) thinks that this leads to the idea of a brain that is secluded from its environment and extracranial body. Kirchhoff admits that there is self-evidencing here too, but synergy, according to (Kirchhoff 2018), provides this connection without appealing to representations. He offers the idea of ‘synergy’ (see Kelso 1995; 2009), thanks to which in a nontrivial way, “organisms can minimize complexity” (Kirchhoff 2018, 250). ‘Synergy,’ he proposes, is “shortly-lived assembly” (Kirchhoff 2018, 250, emphasis added). Synergies are “not as static representational as motor programs” (Riley et al. 2012, 23). Assuming

---

12 Hohwy (2014; 2017) notices that in the PP framework, when one has a hypothesis that explains some of its evidence, it also provides evidence for itself. In his own words:

The internal model that generates hypotheses that over time makes the evidence most likely, and does so most precisely and simply, will have its own evidence maximized. That is, as a model generates hypotheses that explain away occurring surprising evidence (i.e., minimize prediction error) it maximizes the evidence for itself. Prediction error minimization thus constitutes self-evidencing. This is then the doctrine of the self-evidencing brain (Hohwy 2014, 6).

To Hohwy, “this enforces an evidentiary boundary between it and the external causes of sensory input harbored in the environment and in the rest of the body” (Hohwy 2014, 1). This leads to a strict separation between the brain and its environment including the extracranial body, which contradicts the underlying claims of 4E cognition. Thus, self-evidencing in the PP framework creates the need for representations to bridge the gap between these elements.
that this is true, there is no gap between the world, the body, and the agent, which obviates the need for representations to bridge the gap between them. Thus, we need not have representations for the PP hypothesis to be true, and PP is perfectly compatible with 4E cognition. In arguing this, what he has in mind is obviously low-level models. Low-level models are thought to be responsible for short term, on-the-fly, and almost reflexive activities. Even the passage he quotes from Clark reveals this fact: “‘... model’ means in low-level free energy minimization accounts...” (Clark 2016, 14, emphasis added).

In considering these points, he argues against the functional property “decoupleability” of those posits. The representational posits of PP (as I argue, only the ones of low-level models in PP) are not decoupleable since they are in a causal coupling with external states of affairs. These representational posits do not bear any content and function as a merely causal mediator “because internal and external states cause one another in a circular and reciprocal fashion” (Kirchhoff 2018, 251). “They do not seem to imply the presence of inner model or content-bearing states” (Clark 2016, 14). Kirchhoff’s position could point to another interesting conclusion. He seems to understand representations in a more theoretical sense because he does not directly reject the existence of representational posits. He sets out to illustrate how they are causally coupled with external states, and “decoupleability” seems to be his criterion for a representation,13 but this is not explicitly stated in (Kirchhoff 2018).

Another point undermining Kirchhoff’s nonrepresentationality argument relates to the fact that while for Hohwy’s argument, pointing out a single representation proves that cognition involves representations, for Kirchhoff, it does not. He must have shown that not only the generative model but also all other posits of PP are not representations. However, he only discusses the generative model while remaining silent about the other posits listed above, which may be representations. To him, if any of these posits happens to be decoupleable, this invalidates his argument.

Finally, we can ask which version of 4E cognition he refers to. If he argues that certain cognitive phenomena are nonrepresentational and thus

13 Remember, for instance, that Wheeler (2005; 2008) contends that minimal representations do not need to exhibit ‘decoupleability.’
contentless, then he loses sight of a myriad of nonradical versions of 4E cognition. Indeed, only some extremely radical versions of enactivism and embodiment deploy contentless posits such as Hutto and Myin’s (2013). Kirchhoff’s nonrepresentationality thesis claims that “at least some kinds of cognitive activities” (2018, 244, emphasis added) are nonrepresentational and thus contentless. However, proponents of radical 4E cognition make use of contentless posits in explaining cognitive phenomena of various levels. Thus, unless Kirchhoff specifies what “at least some kinds of cognitive activities” are and to what extent he can explain cognitive phenomena with contentless PP architecture, it also appears impossible to determine exactly which versions of 4E cognition Kirchhoff refers to.

In sum, Hohwy’s posits have the representational function of mirroring nature, and their content is organism-salient. Since he takes representations from this representational function, he does not seem to bother contemplating whether they are content-bearing states or not. Starting from the cognitive function of a mind employing PP, Kirchhoff derives the nondecoupleability of representational posits: they are causal mediators. His second argument relates to the so-called content of these posits. For Kirchhoff, surprise being relative to a given model means that an organism does not harbor a model of the environment, but becomes the model of its environment. Accordingly, such posits do not bear any content. Moreover, while Kirchhoff seems to only discuss low-level models, Hohwy deals with the whole hierarchical architecture performing PEM. Finally, in both articles, it is not palpable which versions of 4E cognition are being considered. Possible candidates for Hohwy’s account might only be searched among certain deflated versions, and Kirchhoff’s account can be compatible to radical versions of 4E cognition.

5. Conclusion

In section 2, a selective and brief overview of PP and 4E cognition is provided, and how the question of their relation is tied to the features of representational posits they deploy is demonstrated. In section 3, the general schema of the (non) representationality argument mostly appealed to in arguing for the compatibility of PP and 4E cognition is given, and it is
shown why any instance of such an argument falls short of proving that PP and 4E cognition are compatible. In the last section, Hohwy and Kirchhoff’s arguments are analyzed in detail to illustrate how the criticism raised here is a significant one and the consequences of ignoring it.

The argument presented here is a good one only insofar as it can show that if it depends on either nonrepresentational PP or representational 4E cognition, any argument for the compatibility of PP and 4E cognition can become much more complicated than treated so far in the literature of cognitive science and philosophy of mind. If it is a good one, its significance does not lie in detecting exactly which version of 4E cognition or which kinds of representational posits the authors in question refer to, but in genuinely demonstrating why such arguments are always insufficient in what they argue for.

It is definite that the representational posits of PP are radically different from the older ones in various respects and that the anti-representationalism of 4E cognition is not homogenous. Given this, arguments for nonrepresentational PP or representational 4E cognition fail to appreciate diversity and heterogeneity in this pertinent debate. Rather, an ontological analysis of these representational posits and of their capacities to satisfy skepticism of 4E cognition should be central if one aims to argue for the compatibility of 4E cognition and PP.

To conclude, neither of the theories have a uniform understanding of the debate. When arguing against representations, proponents of 4E cognition have different and distinct ideas of representations in mind. When evaluating PP, its proponents refer to various parts of the whole multilayered structure. This variety hinders the validity of any argument for the compatibility of these theories if they propose instances of the (non) representationality debate sketched in section 3. Hence, without depending on any definition of representation, one must refer to similarities and dissimilarities in the properties of these posits in both theories. With this paper I hope to have shown that the growing tendency in the literature to argue about the existence of representations in these theories to establish their compatibility should be abandoned because it seems quite impossible to resolve the complexities of representationalism debates on PP and 4E cognition in a single argument. For this reason, this paper contends that such
arguments will always be partial and fail to be conclusive. Instead, an effort to reveal the representational properties of the posits of each particular theory and to compare these properties shall better serve this debate.

References


Clark, Andy. 2015a. “Predicting Peace: The End of the Representation Wars-A Reply to Michael Madary.” In *Philosophy and Predictive Processing*, edited by...
https://doi.org/10.15502/9783958570979


https://doi.org/10.1093/acprof:oso/9780190217013.001.0001


https://doi.org/10.1016/j.tics.2009.01.007


https://doi.org/10.7551/mitpress/2524.001.0001


