Philosophical Explorations, 2018
Vol. 21, No. 2, 179–186, https://doi.org/10.1080/13869795.2018.1479440
This paper is part of the special issue "Enactivism, Representationalism, and Predictive Processing" guest edited by Krzysztof Dołęga, Luke Roelofs, and Tobias Schlicht.

Introduction

Krzysztof Dołęga, Luke Roelofs and Tobias Schlicht*

Institute of Philosophy II, Ruhr-Universität Bochum, Bochum, Germany

(*Received 18 May 2018; final version received 18 May 2018*)

The papers in this special issue make important contributions to a longstanding debate about how we should conceive of and explain mental phenomena. In other words, they make a case about the best philosophical paradigm for cognitive science. The two main competing approaches, hotly debated for several decades, are representationalism and enactivism. However, recent developments in disciplines such as machine learning and computational neuroscience have fostered a proliferation of intermediate approaches, leading to the emergence of completely new positions, in particular the Predictive Processing approach. Here, we will consider the different approaches discussed in this volume.

Keywords: enactivism; predictive processing; representationalism; philosophy of mind

Representationalism

Representationalism has been the dominant paradigm in philosophy and cognitive science since the middle of the twentieth century. It is characterised by explaining mental processes in terms of representational states, ones which exhibit the property of being intentionally directed at or being "about" something other than themselves. It posits, that is, that when someone thinks about something, this involves an internal state which is itself about the thing which is the target of their thought. Changes in thinking involve changes in such representational states and their relations to each other. Understanding the mind is thus taken to require understanding putative "states of the nervous system that have content, that refer to concrete or abstract entities (or even fictional entities), properties, and events" (Carey 2009, 5).

Traditionally, such representational states are thought of as being "sandwiched" (Hurley 1998) between sensory inputs and motor outputs, and their transformations are considered to be analysable in abstraction from the details of perception and action. The brain's task is to enable successful action by processing sensory input into cognitive states that represent different aspects of the world. When I look at the coffee mug in front of me, for instance, sensory information hitting my retina is processed in specialised modules that eventually produce a detailed three-dimensional representation of the mug that can guide actions like grasping it (Marr 1982).

*Corresponding author. Email: tobias.schlicht@rub.de

There are several varieties of representationalism. Most famously, the Computational Theory of Mind (Fodor 1975; Pylyshyn 1984) holds that mental phenomena should be explained in terms of *computations* performed over representational states. It draws on the metaphor of the brain being a kind of computer, which manipulates internal sets of symbols much like a digital computer does. By contrast, the parallel distributed processing movement presented a departure from the simple computer metaphor by modelling cognition using connectionist networks inspired by the architecture of the brain (Smolensky 1988; Churchland and Seijnokwski 1992). However, even though such artificial neural networks process input sub-symbolically, most successful models of this kind still rely on the idea that cognition lies in the intermediate stages between perception and action, which store, manipulate, and transform information about external objects before producing the output (O'Brien and Opie 2004).

Enactivism

Enactivism challenges the representationalist paradigm, both in its posit of internal representational states and in its explicit separation of perception, cognition, and action. On this framework, perception, action, and cognition are seen as aspects of the bodily activities of whole organisms (sometimes construed as agents or systems) which can be largely or completely explained without appeal to mental representations (Varela, Thompson, and Rosch 1991; Noë 2004; Chemero 2009; Hutto and Myin 2013; Gallagher 2017). According to Hutto:

> The defining feature of the recent enactivist turn in cognitive science is that it challenges the representationalist paradigm, stressing instead the embedded and embodied nature of cognitive acts. In particular, the movement has been associated with the rejection of the very idea that we can make sense of the basis of everyday skills in terms of the manipulation of underlying tacit representations of a pre-given world. (Hutto 2005, 389)

The brain, on such views, is seen "as part of the body-environment system (not only regulating the body, but regulated by the body and its affective processes) […]" (Gallagher 2017, 24). It is, as Jesse Prinz puts it, "set up to be set off" (2004, 55) by aspects of the environment, in a way moulded by plastic changes resulting from prior experiences – without the need to *represent* those aspects that set it off. Thus, on the enactive framework, cognition is conceived as something which unfolds across the whole of the organism as well as its dynamical responses to the environment. My perception of the coffee mug not only involves multiple *actions* like eye, head, and body movements (gaze turning etc.), but is in the service of detecting action possibilities (like grasping, say) from the very start (Gibson 1979). The *cognition* of the coffee mug is not sandwiched between perception and action, but constituted by the sensitive and skilful way that they intermingle.

Enactivism, like representationalism, is diverse. One family of views, *sensorimotor enactivism*, focuses on the relationship between cognition, perception, and action. Crucially, perception is conceived as something we do, not something that happens to a passive subject of experience (Noë 2004). More specifically, the range of an organism's sensorimotor skills together with the layout of the environment determines the complex set of action possibilities (affordances) specific to this agent (more on affordances below). An important aspect here is the right kind of *coupling* between agent and world which forces us to treat them as one complex cognitive system, whereby a relevant portion of cognitive processing is *outsourced* by the agent *onto the world*.

According to a second brand, *autopoietic enactivism*, cognition emerges specifically from the self-organising and self-producing activities of *living organisms*. Characteristic for these views is the alleged life-mind-continuity thesis, which holds that cognitive structures display the same organisational features as living structures, such that "mind is life-like and life is mind-like" (Thompson 2007, 128ff).

However, the kind of enactivism central to the debate with representationalism is what Hutto and Myin (2013, 2017) call *radical enactivism*. This third variety is defined by a crucial negative claim: that "basic" cognition – encompassing all the cognitive activities of most non-human animals, including perception and action, all the way to forms of imagining and remembering – does not involve the processing of any mental representations. Representation appears only with distinctively human capacities like language, and the social practices that come with it.

Predictive processing

The third and most recent participant in this debate originates in computational neuroscience. Proponents of so-called *Predictive Processing (or PP)* models of the brain claim to have found "the first truly unifying account of perception, cognition and action" (Clark 2016, 2) by conceiving of the brain as a top-down "prediction machine". On this view, cognition consists in the continuous process of testing hypotheses about the incoming sensory stimulation, generated by a hierarchical generative model of the world that is constantly updated based on prediction error signals (Friston 2010; Hohwy 2013; Dołęga 2017). When I perceive the coffee mug, my brain processes have already formed a set of top-down expectations about the incoming sensory information based on the most likely hypothesis given my previous sensory state and my prior knowledge about the context I am in. These expectations are matched against the actual sensory input resulting in bottom-up signalling of deviations from the prediction. On the PP view, the neural channels conveying sensory input are reinterpreted as passing only the prediction error, which is then used to enforce an update of the internal model that issued the top-down expectations in the first place. This process continues until the model (more or less) matches the feedback and can thus count as accurate with respect to the state of the world. Importantly, since the sensory input is usually ambiguous with respect to its cause, the brain is forced to formulate probabilistic hypotheses about what caused the input. Thus, the process of model deployment and revision postulated by PP is meant to serve as an approximation of the optimal updating recipe given by Bayes' rule. Viewed from this perspective, perception, action, attention, and other cognitive phenomena are all ultimately "doing the same thing" (Hohwy 2013, 2): The overarching goal (and unifying principle) of the brain's perceptual and cognitive activities is to minimise prediction error, which can also be understood as information theoretic "surprise".

Given the promise that this Bayesian prediction machine account could yield a unifying framework for cognition, it is not surprising that both representationalists and enactivists have tried to claim it for themselves, resulting in a hot debate regarding whether this framework should be interpreted in representationalist (Hohwy 2013; Clark 2015; Gładziejewski 2015; Gładziejewski and Miłkowski 2017; Kiefer and Hohwy 2017) or enactivist terms (Kirchhoff 2017; Hutto and Myin 2017; Kirchhoff and Robertson 2018). The key question concerns the hierarchical generative models whose predictions are compared to sensory input, generating greater or lesser prediction error. Are these models representations of the external causes of sensations, or is such a designation an unhelpful philosophical gloss that adds nothing to our understanding of how they work?

Action-oriented representations

Some philosophers and cognitive scientists have attempted to achieve a compromise between representationalism and enactivism by modifying the notion of representation, most notably through the posit of "action-oriented representations".

Action-oriented representations can be seen as part of an evolving response to Gibson's (1979), initially anti-representational, notion of *affordances* – possibilities for action that emerge from the coupling between an organism and its environment and are perceptually detected by that organism. On his view, the primary objects of perception are not the objective properties of the environment, but the practical relations between the environment and an organism. For example, a surface may be at the right height, and sufficiently horizontal and rigid, for me to sit on it. According to Gibson, what my perceptual system primarily needs to detect is this "possibility-for-sitting", not the intrinsic properties that underlie it. Another creature which could not sit on it would perceive it differently, even though the surface's intrinsic properties remained unchanged. The organism's bodily and sensorimotor capacities together with the properties of the environment determine the affordances offered by the environment to that particular organism.[1]

Although Gibson originally denied that perceiving affordances involved representations, others disagreed. Millikan, for example, argues:

> Gibsonians have generally assumed that if there were such things as inner representations they would have to be things calculated over, vehicles of inference, and hence, that the perception of affordances does not involve inner representations. But inner processes mediating the perception of and responses to Gibsonian affordances would certainly involve pushmi-pullyu representations, these being far more primitive than the representations Gibsonians reject. (2004, 159)

Millikan here appeals to her notion of a "pushmi-pullyu-representation" (Millikan 1995), which she conceives of as a kind of representation which both *informs* a subject about the state of the world and *prescribes* specific actions based on (or associated with) this information. That is, pushmi-pullyu representations have both a world-to-mind direction of fit and a mind-to-world "direction of fit". They thus contrast with purely perceptual states and beliefs, which have traditionally been considered as having *only* a mind-to-world direction of fit: Their purpose is to represent or convey how the world is, such that they are flawed when they don't match reality. In this case, they need to be amended. They contrast with desires and intentions, which have been conceived as having *only* a world-to-mind direction of fit: They present how we want the world to be, and are thus meant to guide us in changing it to fit our mental state. For example, if my belief that there is coffee in my mug does not accord with the actual state of affairs, then I need to change my belief; if my desire that there is coffee in my mug does not accord with reality, then I need to change the state of affairs (Anscombe 1963, 56; cf. Searle 1983).

Drawing on the idea that pushmi-pullyu representations may be the most evolutionarily basic form of representation, Clark suggests that cognitive science should deflect enactivist criticisms by constructing theories based on "action-oriented representations":

> In representing (…) the environment as such a complex of possibilities, we create inner states that simultaneously describe partial aspects of the world and prescribe possible actions and interventions. (…) they say how the world is and they prescribe a space of adaptive responses. (Clark 1997, 50)

More precisely, action-oriented representations may be characterised by at least the following three features: first, they are *action-specific*, i.e. they are designed to represent the world in terms of possible actions and interactions. Second, they are *egocentric*, i.e. they are geared towards possible actions and interactions that are possible for the specific agent doing the representing. Finally, they are highly *context-dependent* and may lead to different kinds of agent-object-coupling or agent-agent-coupling that are characteristic of non-social and social interaction (Wheeler 2008, 199). In this sense, they are good candidates for providing an analysis of affordances and for achieving a compromise between the traditional cognitivist paradigm and enactivist concerns about the embodied and situated character of perception and cognition.

Enactivism and Predictive Processing Models constitute a strong departure from the initial Computational Theory of Mind, as developed by Fodor, Pylyshyn and others. In order to retain mental representations in cognitive science explanations, integrating insights from embodied cognitive science and enactivism force us to conceive of representations as action-oriented, egocentric, and context-sensitive. Predictive Processing accounts require a further modification of the notion of representation. As Clark puts it:

> Neural representations, should the hierarchical predictive processing account prove correct, encode probability density distributions in the form of a probabilistic generative model, and the flow of inference respects Bayesian principles that balance prior expectations against new sensory evidence. This … is a departure from traditional understandings of internal representation, and one whose full implications have yet to be understood. (Clark 2013, 188)

As Knill and Pouget (2004, 713) also emphasise, the relevant representations that are allegedly processed by the brain in the service of perception, cognition, and action must be conceived of as representing information probabilistically. Thus, which paradigm we choose in order to explain a specific cognitive phenomenon will have important consequences for our overall view of the mind and brain. It may be that aspects of all paradigms are correct and yet, that none of them will be able to explain the whole range of cognitive capacities in human and non-human animals. As the debate over the best interpretation of PP shows (Dołęga 2017), it is possible that many of the disagreements between enactivists and representationalists are due to distinct explanatory interests and different ways of carving up the explanandum – cognition. For example, enactivists' focus on the organism's environmental coupling may be well suited to describe cognition on a larger, coarse-grained scale, which allows for a continuum from adaptive behaviour to sophisticated cognitive skills. On the other hand, representationalists' focus on vehicles and contents of cognition may provide a better framework for finding neural mechanisms which underpin different kinds of flexible behaviours (Boone and Piccinini 2015).

Contributions

The contributions in this volume engage with these debates in different ways and from different sides.

Hutto and Myin, whose book-length formulation of anti-representationalist enactivism is cited in every other paper, open with a response to critics who claim that the disagreement concerning whether mental representations should figure in explanations of cognition is empty, trivial, or merely terminological.

They are followed by two papers defending versions of representationalism or criticising the anti-representationalist arguments used by enactivists. Rupert's paper attempts a

sustained defence of an account of representation he traces back to Dretske (1986), one in which representations – or at least, *mental* representations – get content from their causal-informational relations to external things. This approach has been extensively criticised by enactivists, a critique which Rupert undertakes to answer. Raleigh's paper criticises Hutto and Myin's arguments in favour of a particularly strong anti-representationalist claim, namely that positing internal mental representations is conceptually confused, while being sympathetic towards the more modest claim that many cognitive activities may be best explained in non-representational terms.

The next three papers consider the relationships between representationalism, computationalism, and predictive processing. Kuokkanen and Rusanen's paper argues for the compatibility of enactivism and computationalism. More precisely, they argue that Hutto and Myin's arguments against computationalism fail, even if their related arguments against representationalism succeed. Thus one might reject representational content, and yet still maintain a view of cognition as an essentially computational activity. Kirchhoff and Robertson argue for the compatibility of enactivism and predictive processing, and for the failure of representationalist interpretations of the latter. More specifically, they argue against a particular attempt by Kiefer and Hohwy (2017) to interpret predictive processing as involving measures of *mis*representation, and to that extent as involving representation. Williams' paper focuses on a precursor to modern predictive processing personified by Kenneth Craik and his framework for thinking about representation and cognition modelling the world.

Finally, our last two papers ask what enactivism implies about two specific mental phenomena: emotions and dreams. Hufendiek's paper examines the possibility of understanding emotions as non-representational – as directed at "danger", "insult", "misfortune" but without thereby having representational content. She argues that even when emotions are understood as essentially embodied, the abstractness of what they target (danger, insult, misfortune, etc. – "core relational themes") requires a representational interpretation. Rosen's paper, meanwhile, examines how well enactivism can make sense of dream experience, in which the mind seems to be cut off from all the dynamic interactions with the environment that enactivism appeals to, and yet still seems to display the same richness of structure and quality that waking experience does.

Together, these papers illuminate the central dispute between enactivism and representationalism from a variety of angles. Can we make sense of dreams or emotions in enactivist terms? Can we make sense of predictive processing or computation without positing representational content? Do enactivist arguments against representation in general, or against particular ways of construing representations, succeed? And how much depends on the outcome of these debates? With this special issue we try to make progress on all of these questions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Note

1. It should be noted that the increased interest in the notion of affordances has produced a number of different interpretations of the notion (e.g. Turvey 1992; Reed 1996; Chemero 2009; Rietveld and Kiverstein 2014).

Notes on contributors

Krzysztof Dołęga is a PhD candidate in philosophy at Ruhr Universität Bochum in Germany, where he is supervised by Prof. Tobias Schlicht and Prof. Daniel Dennett. His doctoral project focuses on the place of experience and intentionality in predictive processing, but he is also interested in the framework's explanatory status as well as the general issues relating to the use of computational models in philosophy of mind and cognitive science. His work has been featured in several outlets including *Journal of Consciousness Studies*, *Frontiers in Psychology*, and the 2017 open access collection on the philosophy of predictive processing edited by Prof. Thomas Metzinger and Dr. Wanja Wiese. He is also one of the editors of a forthcoming OUP collected volume on mental representations.

Luke Roelofs is a postdoctoral researcher at Ruhr-Universität Bochum, working in Tobias Schlicht's group on situated social cognition. He works particularly on the interrelations between social cognition, imagination, and empathy, and also on the relationship between consciousness in systems and consciousness in their parts. He has a monograph on the latter topic in production with OUP, and has been published in journals including *Philosophical Studies*, *Erkenntnis*, *Dialectica*, *Dialogue*, *Journal of Consciousness Studies*, *Thought*, *Pacific Philosophical Quarterly*, and others.

Tobias Schlicht is Professor of Philosophy at Ruhr-Universität Bochum, leading a research group on Situated Cognition. His specialisation is topics related to consciousness and cognition. He also writes on Kant's theoretical philosophy and its relevance for contemporary philosophy of mind and cognition, and the philosophical foundations of cognitive science. He has just published an introductory book on Social cognition (Junius 2018) and his work has appeared in *Philosophical Studies*, *Behavioral and Brain Sciences*, *American Philosophical Quarterly*, *Journal of Consciousness*, *Phenomenology and the cognitive sciences, Philosophical Psychology*, *Grazer Philosophische Studien*, *International Journal of Philosophical Studies* and other journals and volumes. He is currently Associate Editor of Philosophical Explorations.

References

Anscombe, G. E. M. 1963. *Intention*. 2nd ed. Oxford: Basil Blackwell. (first edition 1957).

Boone, W., and G. Piccinini. 2015. "The Cognitive Neuroscience Revolution." *Synthese* 193 (5): 1509–1534. doi:10.1007/s11229-015-0783-4.

Carey, S. 2009. *The Origin of Concepts*. Oxford: Oxford University Press.

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

Churchland, P. S., and T. Seijnokwski. 1992. *The Computational Brain*. Cambridge: MIT Press.

Clark, A. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge: MIT Press.

Clark, A. 2013. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioral and Brain Sciences* 36 (3): 181–204.

Clark, A. 2015. "Radical Predictive Processing." *Southern Journal of Philosophy* 53: 3–27.

Clark, A. 2016. *Surfing Uncertainty.* Oxford: Oxford University Press.

Dołęga, K. 2017. "Moderate Predictive Processing." In *Philosophy and Predictive Processing*, edited by T. Metzinger and W. Wiese, 1–19. Frankfurt am Main: Predictive-Mind.net.

Dretske, F. 1986. "Misrepresentation." In *Belief: Form, Content, and Function*, edited by R. Bogdan. Oxford: Oxford University Press.

Fodor, J. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Friston, K. J. 2010. "The Free-energy Principle: A Unified Brain Theory?" *Nature Reviews Neuroscience* 11: 127–138.

Gallagher, S. 2017. *Enactivist Interventions*. Oxford: Oxford University Press.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gładziejewski, P. 2015. "Predictive Coding and Representationalism." *Synthese* 193 (2): 559–582. doi:10.1007/s11229-015-0762-9.

Gładziejewski, P., and M. Miłkowski. 2017. "Structural Representations: Causally Relevant and Different from Detectors." *Biology and Philosophy.* doi:10.1007/s10539-017-9562-6.

Hohwy, J. 2013. *The Predictive Mind*. Oxford: Oxford University Press.

Hurley, S. 1998. *Consciousness in Action*. Cambridge: Cambridge University Press.

Hutto, D. D. 2005. "Knowing What? Radical Versus Conservative Enactivism." *Phenomenology and the Cognitive Sciences* 4 (4): 389–405. doi:10.1007/s11097-005-9001-z.

Hutto, D. D., and E. Myin. 2013. *Radicalizing Enactivism*. Cambridge: MIT Press.

Hutto, D. D., and E. Myin. 2017. *Evolving Enactivism*. Cambridge: MIT Press.
Kiefer, A., and J. Hohwy. 2017. "Content and Misrepresentation in Hierarchical Generative Models." *Synthese*, 1–29. doi:10.1007/s11229-017-1435-7.
Kirchhoff, M. 2017. "Predictive Brains and Embodied, Enactive Cognition: An Introduction to the Special Issue." *Synthese*, 1–12. doi:10.1007/s11229-017-1534-5.
Kirchhoff, M. D., and Robertson, I. 2018. "Enactivism and Predictive Processing: A Non-Representational View." *Philosophical Explorations* 21 (2): 1–21.
Knill, D. C., and Pouget, A. 2004. "The Bayesian Brain: The Role of Uncertainty in Neural Coding and Computation." *Trends in Neuroscience* 27 (12): 712–719.
Marr, S. 1982. *Vision*. Cambridge: MIT Press.
Millikan, R. G. 1995. "Pushmi-pullyu Representations." In *Philosophical Perspectives 9: AI, Connectionism, and Philosophical Psychology*, 185–200. Atascadero.
Millikan, R. G. 2004. *Varieties of Meaning*. Cambridge: MIT Press.
Noë, A. 2004. *Action in Perception*. Cambridge: MIT Press.
O'Brien, G., and J. Opie. 2004. "Notes Toward a Structuralist Theory of Mental Representation." In *Representation in Mind: Approaches to Mental Representation*, edited by H. Clapin, P. Staines, and P. Slezak, 1–20. Oxford: Elsevier.
Prinz, J. 2004. *Gut Reactions*. New York: Oxford University Press.
Pylyshyn, Z. 1984. *Computation and Cognition*. Cambridge: MIT Press.
Reed, E. S. 1996. *Encountering the World. Toward an Ecological Psychology.* Oxford: Oxford University Press.
Rietveld, E., and Kiverstein, J. 2014. "A Rich Landscape of Affordances." *Ecological Psychology* 26 (4): 325–352.
Searle, J. R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: MIT Press.
Smolensky, P. 1988. "On the Proper Treatment of Connectionism." *Behavioral and Brain Sciences* 11: 1–74.
Thompson, E. 2007. *Mind in Life*. Cambridge, Mass: Harvard University Press.
Turvey, M. T. 1992. "Affordances and Prospective Control: An Outline of the Ontology." *Ecological Psychology* 4 (3): 173–187.
Varela, F., E. Thompson, and E. Rosch. 1991. *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
Wheeler, M. 2008. "Minimal Representing: A Response to Gallagher." *International Journal of Philosophical Studies* 16 (3): 371–376.