

Computation and Representation in Cognitive Neuroscience

Gualtiero Piccinini¹

Published online: 27 February 2018

© Springer Science+Business Media B.V., part of Springer Nature 2018

Cognitive neuroscientists explain cognitive capacities in terms of neural computations over neural representations (e.g., Bechtel 2008). By many measures, their explanations are successful. They are so successful that mainstream cognitive psychology and cognitive science are being absorbed within cognitive neuroscience (Boone and Piccinini 2016). If successful scientific explanation is the measure of what's real, then cognition involves neural computation over neural representations.

Some philosophers beg to differ. On one hand, some insist that computational and representational explanations—or, at any rate, computational and representational explanations of a non-neural sort—are distinct and autonomous from neuroscientific ones (Fodor 1997; Burge 2010). Knoll (this issue) updates this autonomist view for the era of cognitive neuroscience. He concedes that neuroscientific evidence can inform psychological explanation. Nevertheless, he defends the classic autonomist view that some representations and computations have causal powers that their neural realizers lack.

On the other hand, antirealists about computation and representation promote non-computational or non-representational explanations of cognition. According to antirealism, computation and representation are at best helpful glosses and at worst misleading metaphors. Cognition is best explained without positing computation and representation.

The impetus for the papers collected here came from a session on Computation and Representation in Cognitive Neuroscience that was held in Ferrara, Italy as part of IACAP 16, Meeting of the International Association for Computing and Philosophy. All the papers in this special issue were submitted in response to a public call for papers. Thanks to Vivian De La Cruz, Alex Morgan, and Eric Thomson for comments on a previous draft. Many thanks to Mariarosaria Taddeo for inviting me to guest edit this special issue and for arranging the blind refereeing of papers when I was a co-author.

✉ Gualtiero Piccinini
piccininig@umsl.edu

¹ St. Louis, USA

One of the most scientifically serious alternatives to computational and representational approaches is ecological psychology (Gibson 1966). Ecological psychologists argue that cognition is primarily explained in terms of dynamical variables characterizing the interaction between agents and environments. According to them, uncovering inner mechanisms is unnecessary. In any case, if we were to look at inner mechanisms, we should conclude that agents do not perform computations over vehicles carrying information. Instead, agents “resonate” with information in the environment.

What does this “resonance” amount to? Resonance has a precise sense in physics, where it means that an external force, which vibrates at specific frequencies, drives a system to increase its oscillatory amplitude. This is probably too specific a notion of resonance for ecological psychology, despite their insistence that psychology is akin to physics. So ecological psychologists owe us an account of resonance.

Raja (this issue) provides an explicit account of resonance for ecological psychology. He proposes that resonance amounts to the coupling of the dynamical system internal to the agent—most relevantly, the central nervous system—to the dynamical system formed by the agent and its environment. For any behavior, ecological psychology posits dynamical variables that characterize the main invariants involved in that behavior. Raja proposes that, for resonance to obtain, the same dynamical variables must capture both external (agent-environment) and internal (central nervous system) dynamics. Raja offers empirical evidence that this coupling between external and internal variables actually occurs.

Raja considers whether resonance is compatible with computation over representations.¹ Yet in the end he, along with most ecological psychologists, proposes resonance as an *alternative* to computation and representation. How is it an alternative? What is non-computational or non-representational about resonance? To answer this question, we need to explicate computation and representation. In recent years, much progress has been made in this regard.

With respect to computation, the most recent and detailed account on offer is that physical computation is a type of mechanistic process. Specifically, computation is the processing of medium independent vehicles by a functional mechanism in accordance with a rule. A vehicle is medium independent just in case it is defined solely in terms of the manipulation of certain degrees of freedom.²

To illustrate, consider a computer programmed to alphabetize a list of words. The computer takes a random list of words as input and produces a list of words in alphabetical order; this input–output relation is the rule the computer follows. The words are strings of states of finitely many types; these are the degrees of freedom

¹ He writes: “there seems to be no reason, in principle, for an ecological explanation and a computational explanation to be incompatible. Actually, they might be seen as two explanations of the same phenomenon at two different scales: the ecological explanation accounts for the changes of the agent-environment system at the agent-environment scale, while the computational explanation unveils the mechanism that makes those changes possible.”

² This is my way of putting it (Piccinini 2007, 2015); I believe it captures the gist of mechanistic accounts even though others might put it slightly differently (Kaplan 2011; Miłkowski 2013; Fresco 2014; Duwell 2017).

the computer processes. The processing is performed by the computer's components: processor, memory, input devices, and output devices. Thus, the computer is a mechanism whose function is manipulating certain degrees of freedom in accordance with a rule.

This mechanistic account has several advantages, two of which are most relevant here. First, the mechanistic account has the right degree of generality to cover all and only the kinds of system that are normally counted as computing. That includes not only digital but also analog and other unconventional types of computation. Second, the mechanistic account does not require that computational vehicles be representations. This allows computation to explain cognition within both representationalist and anti-representationalist frameworks (cf. Orlandi 2014; Villalobos and Dewhurst 2017). Given the mechanistic account, whether neural states are representations is a logically independent question from whether neural processes are computations.

With respect to representation, there is an emerging consensus that the best way to understand representation in the context of cognitive explanation is structural.³ In the present context, structural representation includes four elements: (i) a homomorphism (partial isomorphism) between a system of internal states and their target, (ii) a causal connection from the target to the internal states, (iii) the possibility for the internal states to be decoupled from their target, and (iv) a role in action control. In other words, in order to be a structural representation, a state must belong to a system of states that bear a second-order similarity to their targets, the targets must cause the states to occur but the states can also occur without their target being present, and the states must guide action based on their similarity to their target. When a system's internal states satisfy these conditions, they qualify as representations in a robust sense, which possess semantic content by the lights of a naturalistic theory of semantic content (Neander 2017).

In light of this background, we can answer a question we posed earlier: there is no conflict between resonance and computation over representation. On the contrary, computation over representation provides a mechanistic explanation of resonance. How does the central nervous system come to resonate with environmental information? How is it that the same dynamical variable captures both external (agent-environment) and internal (central nervous system) dynamics? A plausible answer is that the central nervous system receives information from the environment and, to solve the task at hand, processes such information to extract the relevant invariant from it. This shows that resonance—and hence ecological psychology minus its anti-computationalist and anti-representationalist dogma—is compatible with a computational and representational framework (cf. Scarantino 2003). What remains to establish is whether neural processes are actually computations over representations. According to mainstream neuroscience, they are.

In my opinion, there are two general reasons why neural processes are computations (Piccinini and Bahar 2013). First, the variables that are functionally most relevant to neural processing are spike frequency and timing; both appear to be

³ Here is a literature sample: Swoyer (1991), Grush (2004), O'Brien and Opie (2004), Ramsey (2007), Isaac (2013), Shea (2014) and Gładziejewski and Miłkowski (2017).

medium independent. That is to say, what matters most to neural processes is the frequency and timing of neuronal signals rather than anything more specific about spike biophysics. Therefore, the same sort of functional dependencies between signals—the same sort of computation—could be produced by other physical systems so long as such systems process the right sort of signals in the right sort of ways. Second, neural processes respond to information transmitted from physically different environmental sources. In order for information from these different sources to be integrated and processed, it must be transduced into a system of internal states that is neutral between physically different sources—that is, medium-independent vehicles. Processing medium-independent vehicles (in accordance with a rule) is what computation in the general sense amounts to. Therefore, neural processes are computations.

As to representation, there are many reasons to conclude that neural states are structural representations in the structural sense I sketched above. Shagrir (this issue) points out that computational neuroscientists assume that neural states constitute a model of the environment (which may include the body). That is, they assume that neural states are homomorphic to external variables [condition (i)] and are used by the nervous system to guide behavior [condition (iv)]. Shagrir argues that this representational assumption helps computational neuroscientists discover neural functions and explain why specific neural computations are appropriate to the task.

Maley (this issue) defends a similar conclusion from a different angle. He points out that the main variables involved in neural processes—especially spike frequency and timing, but possibly other variables as well—vary monotonically as a function of their external causes. This conforms to condition (ii) above and leads to a homomorphism between the internal variables and their target [condition (i)], which in turn allows these variables to guide behavior [condition (iv)]. Maley points out that this monotonic variation of neural representations relative to their targets makes them different from digital representations and similar to the kind of representation employed by analog computers. He concludes that, therefore, brains are a kind of analog computer.⁴

Plebe and De La Cruz (this issue) make a rigorous and thorough case that neural states are structural representations. First, they formally define a homomorphism between neural states and their target. Then they argue explicitly that such a homomorphism is causally driven by its target, can be decoupled from its target, and guides action. Therefore, neural states satisfy all four conditions for being structural representations.

Morgan and Piccinini (this issue) situate neural representations, understood as structural representations, within the debate about the nature of intentionality. They address naturalistic theories of intentionality that hope to account for intentionality in terms of mental representation. They argue that such theories should look at neural representations as the underpinning of intentionality. Yet they also argue that

⁴ Appearances notwithstanding, Maley's conclusion is consistent with Piccinini and Bahar's (2013) conclusion that neural computation is neither digital nor analog but *sui generis*. This is because Piccinini and Bahar define analog computers independently of the kind of representation they typically employ.

simply appealing to neural representations—structural as they may be—is not enough to fully explain intentionality. To explain intentionality, a more specific kind of neural representation needs to be identified.

A theoretical framework that has attracted a lot of attention in recent years posits that the brain is a Bayesian inference machine. One way to spell this out is to posit that brains perform Bayesian inferences by building generative models of their environment's causal structure and then striving to minimize such models' prediction errors. This approach goes by the name of predictive processing. Expanding on Clark (2015), Williams (this issue) argues that the kind of generative models posited by predictive processing are structural representations in a robust sense. Yet such structural representations are oriented towards the specific needs and capacities of specific organisms in a way that makes them especially suited to overcome some limitations of more traditional approaches to mental representation.

Anti-representationalists have their easiest cases with real-time sensorimotor engagement with the world—when the action is right in front of you, it's somewhat plausible that you could explain what's happening without invoking internal representations. By contrast, anti-representationalists have their hardest cases when there is mental activity without any direct sensorimotor engagement with the world. Examples include dreaming, extreme paralysis such as locked-in syndrome, and minimally conscious state (MCS).

Noh (this issue) takes a close look at the last of these examples. MCS is a neural condition that predicts recovery of some cognitive function. MCS is most usefully contrasted with the vegetative state, which predicts lack of recovery. The standard clinical method for distinguishing between vegetative state and MCS is to observe signs of intentional behavior such as the ability to follow commands. If such ability is manifested, patients are diagnosed as MCS. Otherwise, patients are diagnosed as vegetative. Recent advances in neuroimaging have shown that some patients that are traditionally diagnosed as vegetative may warrant a MCS diagnosis instead. Noh provides a detailed analysis of the reliability of the inference involved in attributing patients intentional behavior or the ability to answer questions solely on the basis of their neural activity, without any overt behavior. The methods employed by neurologists rely on the representational content of the patients' neural states. Noh shows that such methods work.

The defenses of neural representation listed so far are primarily based on inference, modeling, and other theoretical considerations. While they often appeal to empirical evidence, they are at least consistent with the standard assumption that representations are theoretical posits to be confirmed or disconfirmed based on their role within a theory of cognition. Thomson and Piccinini (this issue) challenge that assumption. They point out that since neuroscientists began to posit representations in the nineteenth century—well before the beginning of the current representation wars—representations have become observable. Experimental neuroscientists have developed multiple techniques to observe and manipulate representations experimentally. As a result, there is a great deal of direct empirical evidence for neural representations—including structural representations in the sense defined above.

References

- Bechtel, W. (2008). *Mental mechanisms*. New York: Taylor and Francis.
- Boone, W., & Piccinini, G. (2016). The cognitive neuroscience revolution. *Synthese*, 193(5), 1509–1534.
- Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press.
- Clark, A. (2015). *Surfing uncertainty*. London: Oxford University Press.
- Duwell, A. (2017). Exploring the frontiers of computation: Measurement based quantum computers and the mechanistic view of computation. In A. Bokulich & J. Floyd (Eds.), *Turing 100: Philosophical explorations of the legacy of Alan Turing. Boston studies in the philosophy and history of science* (Vol. 324, pp. 219–232). Berlin: Springer.
- Fodor, J. (1997). Special sciences: Still autonomous after all these years. *Nous*, 31, 149–163.
- Fresco, N. (2014). *Physical computation and cognitive science*. Berlin: Springer.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.
- Gładziejewski, P., & Miłkowski, M. (2017). Structural representations: Causally relevant and different from detectors. *Biology and Philosophy*, 32, 337–355.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Science*, 27, 377–442.
- Isaac, A. M. (2013). Objective similarity and mental representation. *Australasian Journal of Philosophy*, 91, 683–704.
- Kaplan, D. M. (2011). Explanation and description in computational neuroscience. *Synthese*, 183(3), 339–373.
- Miłkowski, M. (2013). *Explaining the computational mind*. Cambridge, MA: MIT Press.
- Neander, K. (2017). *A mark of the mental: In defense of informational teleosemantics*. Cambridge, MA: MIT Press.
- O'Brien, G., & Opie, J. (2004). Notes toward a structuralist theory of mental representation. In H. Clapin, P. Staines, & P. Slezak (Eds.), *Representation in mind—New approaches to mental representation*. Amsterdam: Elsevier.
- Orlandi, N. (2014). *The innocent eye: Why vision is not a cognitive process*. New York: Oxford University Press.
- Piccinini, G. (2007). Computing mechanisms. *Philosophy of Science*, 74(4), 501–526.
- Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford: Oxford University Press.
- Piccinini, G., & Bahar, S. (2013). Neural computation and the computational theory of cognition. *Cognitive Science*, 34, 453–488.
- Ramsey, W. M. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.
- Scarantino, A. (2003). Affordances explained. *Philosophy of Science*, 70, 949–961.
- Shea, N. (2014). Exploitable isomorphism and structural representation. *Proceedings of the Aristotelian Society*, 114, 123–144.
- Swoyer, C. (1991). Structural representation and surrogate reasoning. *Synthese*, 87, 449–508.
- Villalobos, M., & Dewhurst, J. (2017). Enactive autonomy in computational systems. *Synthese*. <https://doi.org/10.1007/s11229-017-1386-z>.