



“An incomplete theory of the mind”

Javier Bernacer*, Jose Ignacio Murillo

Mind-Brain Group, Institute for Culture and Society, Universidad de Navarra, Pamplona, Spain

*Correspondence: javier.bernacer@gmail.com

Edited by:

Shimon Edelman, Cornell University, USA

Reviewed by:

Axel Cleeremans, Université Libre de Bruxelles, Belgium

A commentary on

“Whatever next? Predictive brains, situated agents, and the future of cognitive science” by Clark, A. (*in press*). *Behav. Brain Sci.*

Clark offers a detailed description of certain current views of the brain that define it as a “prediction machine.” According to these accounts, it is a hierarchically organized processing unit whose main role is elaborate top-down predictions about sensory inputs, and compute a bottom-up prediction error (PE) which will help refine future predictions. The ultimate stable state, says Clark citing Mumford (1992), would be the delivery of a signal from the cerebral cortex to lower areas which would fully predict the sensory information they are receiving. In this state, the PE signal would not exist at all. In Clark’s opinion, the brain is a “black box” which receives information from the world without a direct access to it: “[the brain] must discover information about the likely causes of impinging signals without any form of direct access to their source”; that is, it just perceives the perturbations that these signals are causing on its own state, makes predictions about the nature of the signals, and elaborates a response.

In our opinion, the main problem of this view is the description of cognition as prediction, and prediction as some sort of representation. Perception, hence, is just a mean to correct previous predictions. This eludes the question about the very nature of cognition. In fact, if we understand cognition as a representative activity, there are two possible interpretations: cognition could be either (1) the process that generates the representation or (2) the resulting representation itself. In (1) we identify cognition with a causal process, and this neglects the fact that a representative activity is only possible if guided by a previous cognition. In (2) we face the problem of explaining what is to

know the representation: if cognition is a representation, we need a representation to know the previous one, and so on *in infinitum*. However, inasmuch as cognition goes beyond the mere generation of a prediction, a representation can be compared with the reality that represents. This comparison is a cognitive activity irreducible to the representation itself.

In addition, *representation* is not possible without some kind of previous “presentation.” Clark’s example of the black box is very illustrative. He overlooks that cognition itself means the suppression of the “black box problem.” The confinement of an entity into a “black box” is a problem only if that entity is a cognitive system; that is, only a unit with the ability of cognition can be deprived of the possibility of comparing a representation with the external reality that represents. We think this clarification is especially important to understand the more basic sensory and perceptual levels. Vision, for instance, is not a representation of the world, but a codification of some information about the world. This intrinsic relation between cognition and the world can be called *intentionality*, using a word reintroduced by phenomenology, but only if, unlike phenomenology (Gallagher and Zahavi, 2008), we disentangle intentionality from consciousness. Following the example, vision does not need to be conscious to be cognition; it just need to apprehend some information (i.e., color) that is in relation to the world. This could be considered a “presentation” (in a weak sense) of reality.

This understanding of the cognitive activity is necessary to accept the existence of representations, which are an important way of knowledge, although they are not the first level. Assuming this, the interaction between the brain and the world has not only an effective or impinging character, as Clark states. The basic and more general level of cognition consists of the abstraction

or separation of “information” by sensory systems from the effective causality which transports it (Polo, 2002). This process generates information as such, and makes it available to superior levels of cognition and to behavior. In our opinion, such aspect of cognition is not explained by the Bayesian modeling of the brain.

One of the consequences of this omission is that the unified science of mind and action Clark proposes makes a deficient distinction between cognition and action. Action is a causal intervention of the living being in the world, but it is only an action inasmuch as it is informed by cognition. This leads to a different, although connected, sense of intentionality (Moya, 1990). The dissolution of cognition in a causal effective process blurs, as a consequence, the distinctive features of actions.

However, the representational model proposed by Clark could be useful to understand some sort of secondary levels of cognition and practical action. In fact, we would like to propose a neural substrate for the so-called “Bayesian brain.” This should be a brain region able to integrate the top-down prediction with the bottom-up PE. Clark cautiously proposes a possible candidate for such interaction between predictions and PEs citing Friston (2005, 2009) and Mumford (1992) models. However, we believe a different option should be taken into account.

The striatum is the main receptor of cognitive, sensory, motor, and emotional information within the basal ganglia, a group of nuclei intimately involved in action from a wide point of view (decision-making, action selection, procedural memory, and instrumental learning, among others; Redgrave and Gurney, 2006). This special involvement in action puts the striatum in a preferential position to be a biological substrate of the “Bayesian brain.” On the other hand, midbrain dopaminergic neurons located in

the substantia nigra pars compacta and ventral tegmental area have been reported to convey a PE signal to striatal areas (Schultz, 1998). Its representation in the striatum has been widely demonstrated in humans (Tobler et al., 2006). In fact, Clark mentions the role of this nigrostriatal PE signal in different sections of his article. Moreover, anatomical studies have also demonstrated that the same striatal neurons targeted by a dopaminergic input also receive projections in the same synaptic spine from associative, motor, and limbic cortices – those possibly carrying the predictions Clark mentions in his model (Gerfen and Surmeier, 2011). The result of this interaction between the neural signals of priors (from the cortex) and errors (from the midbrain), together with the modulation from striatal interneurons (Bernacer et al., 2012), could determine the activity of striatal projection neurons, thus facilitating or inhibiting the performance of an action.

In conclusion, a holistic theory of the mind should distinguish cognition from action in a sharper way than Clark does. Action is a causal process arising in the subject and aimed to the outside world, but it

is also guided by cognition itself. This is the key feature of action. We also believe there are enough evidences to propose the striatum, part of the basal ganglia, as the neural substrate of Clark's proposal on practical action. On the other hand, the difficulty of disentangling information from causality in laboratory experiments makes difficult to find a neural substrate for sensory cognition.

ACKNOWLEDGMENT

We are grateful to all members of the Mind-Brain Group (Institute for Culture and Society) for their helpful comments about this matter.

REFERENCES

- Bernacer, J., Prensa, L., and Gimenez-Amaya, J. M. (2012). Distribution of GABAergic interneurons and dopaminergic cells in the functional territories of the human striatum. *PLoS ONE* 7, e30504. doi: 10.1371/journal.pone.0030504
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836.
- Friston, K. (2009). The free energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301.
- Gallagher, S., and Zahavi, D. (2008). *The Phenomenological Mind. An Introduction to Philosophy of Mind and Cognitive Science*. London and New York: Routledge.
- Gerfen, C. R., and Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* 34, 441–466.
- Moya, C. J. (1990). *The Philosophy Of Action. An Introduction*. Cambridge: Polity Press.
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* 66, 241–251.
- Polo, L. (2002). La cibernética como lógica de la vida. *Studia Poliana* 4, 9–17.
- Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Tobler, P. N., O'Doherty, J. P., Dolan, R. J., and Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *J. Neurophysiol.* 95, 301–310.

Received: 30 August 2012; accepted: 30 September 2012; published online: 19 October 2012.

Citation: Bernacer J and Murillo JI (2012) "An incomplete theory of the mind". *Front. Psychology* 3:418. doi: 10.3389/fpsyg.2012.00418

This article was submitted to *Frontiers in Theoretical and Philosophical Psychology*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Bernacer and Murillo. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.