



The reward of unification: A realist reading of the predictive processing theory



Majid Davoody Beni

Amirkabir University of Technology, Philosophy of Science Group, Department of Management, Science and Technology, No. 424, Hafez Street, Tehran, PO Box 3313–15875, Iran

A B S T R A C T

In a recent paper, Colombo and Wright (C & W for short) built on their neat assessment of the activity of mesocorticolimbic dopaminergic (DA) systems to argue that hierarchical predictive processing theory of the brain (PTB) is not the grand unifying theory that has been claimed by its advocates. To the contrary, they argued that the scientific practice is consistent with a pluralist reading (Colombo & Wright, 2016). Despite its reliance on solid experimental resources, C & W's defence of explanatory pluralism is adorned with sophisticated philosophical assumptions. Of course, this provides all the more reason to admire C & W's enterprise on account of its tasteful combination of scientific and philosophical insights. However, despite being conscious about the relation of their works to the philosophy of science, C & W have presented their pluralistic construal as if it is a direct consequence of the scientific status quo.

This paper aims to defend a unificationist account. C & W's endeavour might successfully block the way of some discredited forms of unificationism (such as micro-reductionism and Nagelian model). However, in this paper, I argue that it fails to establish its pluralist reading in the face of the more advanced forms of unificationism that are at issue in the contemporary philosophy of science (mechanistic accounts or structural realist accounts). After introducing these more sophisticated theories of unification schematically, I refer to the scientific practice to show how the free-energy formulation of PTB (as well as a case that has been mentioned by C & W as evidence for pluralistic reading) could be construed along the lines of unificationism. I end the paper with reminding that the reward of unification is realism.

1. Predictive processing and the unificationist reading

In a nutshell, PTB holds that the brain is a Probabilistic-Bayesian predictive machine that aims at minimising its sensorimotor-perceptual error as regards the causal structure of the world (Huang and Rao 2011). The expectations about the causal structure of the world are encoded into the backward connection, whereas prediction errors are transmitted through the feedback loops to the higher levels of cognition. PTB is formulated along the lines of free-energy principle which indicates that biological self-organising systems aim at minimising long-term average energy (Friston, Daunizeau, Kilner, & Kiebel, 2010). PTB has been praised on account of its great unifying virtue and as a grand unifying theory. C & W dubbed this interpretation the grand-unifying-theory reading (GUTR for short). C & W also referred to different scientific and philosophical interpretations of PTB to indicate that GUTR assumes that the theory is maximally explanatory, deeply unifying, and singularly fundamental. Let me emphasise the fact that these conditions are extracted out of diverse contexts by C & W.

Moreover, the discussion is inlaid with technical philosophical elements. C & W remarked that because “it is an open empirical question whether, and how, PTB relates to other theories and hypotheses, this question should be answered on case-by-case grounds in light of both precise explications of concepts like UNIFICATION, REDUCTION, and EXPLANATION, as well as actual scientific practice” (2). This proposal opens the door to philosophical discussions, and indeed several advocates of the so-called GUTR are notable philosophers (Clark 2013a, 2013b, 2016; Hohwy 2014; Gładziejewski, 2016). Before engaging with evaluation of C & W's philosophical argumentation against GUTR, I briefly overview the scientific evidence that they propose in order to support their arguments.

2. Scientific evidence against GUTR

Part of the support that GUTR receives from the scientific field roots in the fact that the free-energy formulation of PTB could successfully assimilate the diverse theories of the role of dopamine in its Bayesian

E-mail address: m.davoody@auta.c.ir.

framework. C & W, on the other hand, argued that since theories concerning the activity of mesocorticolimbic dopaminergic (DA) systems could not be absorbed into the framework of PTB without significant *explanatory loss*, the unificationist reading flaws and explanatory pluralism looms large. Their main argument for establishing their pluralistic conclusion is that: “if pluralism were correct, then the scientific investigation of DA activity would demand multiple, diverse epistemic tools without a requirement to collapse into a fundamental theory of how brains work. As this multiplicity and diversity are just what is observed in current scientific practice, pluralism is vindicated.” (Colombo & Wright, 2016, p. 2). Let us attend to C & W’s account of the diverse epistemic tools of the mentioned investigations.

Dopamine (DA) is a catecholaminergic neurotransmitter realised by DA neurones, whose axons are spread through numerous cortical and subcortical brain areas. The hypotheses of anhedonia (HED), incentive salience (IS), and reward prediction error (RPE) are among theories that offer to account for the function of DA. These hypotheses receive confirmation from different kinds of evidence from humans and other animals, but none of them provides a comprehensive explanation of DA complexities on the basis of a single set of fundamental principles (Colombo & Wright, 2016, p. 7). This is because the role of DA has been accounted for through different biomedical theories. HED concerns localising the hedonic ‘hotspots’ in the VP and rostromedial shell of the nucleus accumbens (NAc) so that the complex interaction between DA operations and opioid systems could be used in the account of the brains’ ability to process pleasure (Pecina & Berridge, 2005). IS theory, on the other hand, indicates that the release of DA by mesencephalic structures like the VTA encodes “incentive” value to objects or events (Berridge and Robinson, 1998). The mechanisms that are modelled in HED and IS theories are different not least because IS attributions need not be conscious or involve feelings of pleasure (Colombo & Wright, 2016, p. 6). RPE is yet another theory that aims at setting a connection between the patterns of DA activation on the one hand, and a computational signal called reward prediction error on the other hand. RPE concerns the differences between the expected and actual experienced magnitudes of reward and drives decision-formation and learning for different families of reinforcement-learning algorithms (Sutton & Barto, 1998).

According to C & W, these theories provide different partial models of DA, and the scientific communities rely on these models for different explanatory purposes. Notice that C & W do not argue that the three mentioned theories could not be unified by being absorbed into the information-theoretic, probabilistic free-energy formulation of PTB. RPE is the mathematically more precise of the three theory (to which the other two could be connected formally), and it (i.e. ERP) could be easily formulated in terms of the free-energy principle (Friston et al., 2014). However, to derive their pluralistic conclusion, C & W remarked that even if such unification could take place, the assimilating theory will not be *maximally explanatory*. It was for this purpose that C & W emphasised the significance of the explanatory scope of PTB. It is in this vein that they argued that none of the mentioned theories could be intertheoretically reduced to or subsumed under PTB without explanatory loss. That is to say, GUTR-advocates’ attempt at absorbing these theories into the unifying free-energy-formulation of PTB (Friston, Samothrakis, & Montague, 2012) would inevitably result in the elimination of theoretical terms such as “reward” or “value” or their reduction to sensory states. The same will happen to the exchange of the SI-based notion of salience with the PTB-based notion of precision. What is lost in this unifying attempts is the semantic (explanatory) equivalence of the assimilated and assimilating concepts.

Let me recap. Despite the intervention of philosophical notions of explanation and reduction, at times C & W presented their pluralist approach as if it is a straightforward consequence of the practice of neuroscientists. According, to this view, neuroscientists are “led to ask different questions about DA, and to formulate different predictions that are subsequently tested and assessed in a variety of ways”

(Colombo & Wright, 2016, p. 9). However, what may ground pluralism is actually C & W’s insistence on the indispensability of the maximal explanatory scope to the unifying theory. As the assimilating theory lacks this vast scope, unification is condemned to failure. So, philosophical considerations matter.

3. Philosophical argument against GUTR

As I remarked in the previous section, C & W’s reference to scientific practice does not *ipso facto* demonstrate that pluralism is preferable to GUTR. Although the critical value of the C & W’s pointed remarks as regards the lack of philosophical clarity of some of the statements of GUTR can hardly be exaggerated, yet their own statement of the involved philosophical notions, e.g., unification, reduction, etc., is barely either charitable or convincing. Let me elaborate.

C & W declared that PTB could be associated with GUTR only if PTB entails explanatory unification, monism, and reductionism. As I remarked, these conditions are extracted out of diverse philosophical and scientific contexts, rather than being claimed by any specific advocate of PTB. I argue that the unificationist does not need to comply with any of these conditions, at least not in accordance with C & W’s conception of them.

C & W relied on Kitcher’s idea of explanatory unification to suggest that “[u]nificationism names the thesis that explanations are derivations that unify as many descriptions of target phenomena to be explained, u_1, \dots, u_n , from as few stringent argument patterns as possible” (Kitcher, 1981; Colombo & Wright, 2016, p. 3). The unificationist may want to reply that the concept of explanation, when important for the unification purpose at all, may be different to Kitcher’s definition. Let me remind that Kitcher’s view on explanatory unification was barely mentioned by the advocates of GUTR. The issue of explanation found its way to PTB discussions, mainly through Hohwy’s (Hohwy, 2014, p. 146) well-posed articulation, and in the context of the certain philosophical discussion (that concerns Hempel’s problem of self-evidencing and the topic of Inference to the Best Explanation, IBE). So, the advocate of GUTR does not need to commit herself to this account of explanation at all. The next condition, i.e., monism, has been characterised in terms of the uniqueness of the adequate explanation. The unificationist could reply that it is the prospect of finding the unique adequate explanation in the long run, rather than its actual existence, that fuels the unificationist’s quest.

Also, C & W declared that the reductionism—as a condition of unification—has to be defined in terms of epistemic reduction. This formulation puts an emphasis on a hierarchical picture of theories in which the higher-level explanations in psychology play a heuristic role in developing lower-level explanations in cellular and molecular neuroscience, to which the higher level explanations have to be reduced. C & W connoted that this conception of reduction has to be understood in terms of New Wave reductionism as a variety of Nagelian model (see Bickle, 1996; Endicott, 1998), but I suspect that the notion of reduction that C & W referred to lines up with Oppenheim and Putnam’s micro-reductionism. Be that as it may, as I will argue in the next section, the unificationist does not need to commit herself to either micro-reductionism or the Nagelian model. Perhaps it is worth noting that C & W toyed with the idea that the notion of reduction that is at issue in GUTR is not the explanatory reduction so much as a flexible model of reduction that allows for the diversity of epistemic and explanatory tools. They even allude to the possibility of seeing the intended form of explanation afforded by PTB not as reductionistic so much as mechanistic (4). However, they waved away this rather significant point too hastily, by remarking that “until advocates do the work necessary to demonstrate genuine intertheoretic reductions, rather than just suggestively assert them, their GUT intuition is unwarranted” (Endicott, 1998). This statement has not been justified in the paper. In the next section, I proceed to hint at some sophisticated versions of unification that do not give way to C & W’s conditions and I will consider the issues of

Download English Version:

<https://daneshyari.com/en/article/6811015>

Download Persian Version:

<https://daneshyari.com/article/6811015>

[Daneshyari.com](https://daneshyari.com)