Studies in History and Philosophy of Biological and Biomedical Sciences 45 (2014) 57–67

Contents lists available at ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

journal homepage: www.elsevier.com/locate/shpsc

Deep and beautiful. The reward prediction error hypothesis of dopamine

Matteo Colombo

Tilburg Center for Logic, General Ethics, and Philosophy of Science, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

ARTICLE INFO

Article history:
Received 12 February 2013
Received in revised form 21 October 2013
Available online 16 November 2013

Keywords:
Dopamine
Reward-prediction error
Explanatory depth
Incentive salience
Reinforcement learning

ABSTRACT

According to the reward-prediction error hypothesis (RPEH) of dopamine, the phasic activity of dopaminergic neurons in the midbrain signals a discrepancy between the predicted and currently experienced reward of a particular event. It can be claimed that this hypothesis is deep, elegant and beautiful, representing one of the largest successes of computational neuroscience. This paper examines this claim, making two contributions to existing literature. First, it draws a comprehensive historical account of the main steps that led to the formulation and subsequent success of the RPEH. Second, in light of this historical account, it explains in which sense the RPEH is explanatory and under which conditions it can be justifiably deemed deeper than the incentive salience hypothesis of dopamine, which is arguably the most prominent contemporary alternative to the RPEH.

1. Introduction

According to the reward-prediction error hypothesis of dopamine (RPEH), the phasic activity of dopaminergic neurons in specific regions in the midbrain signals a discrepancy between the predicted and currently experienced reward of a particular event. The RPEH is widely regarded as one of the largest successes of computational neuroscience. Terrence Sejnowski, a pioneer in computational neuroscience and prominent cognitive scientist, pointed at the RPEH, when, in 2012, he was invited by the online magazine *Edge.org* to answer the question "What is your favorite deep, elegant, or beautiful explanation?" Several researchers in cognitive and brain sciences would agree that this hypothesis "has become the standard model [for explaining dopaminergic activity and reward-based learning] within neuroscience" (Caplin & Dean, 2008, p. 663). Even among critics, the "stunning elegance" and the "beautiful rigor" of the RPEH are recognized (Berridge, 2007, pp. 399, 403).

However, the type of information coded by dopaminergic transmission—along with its functional role in cognition and behaviour—is very likely to go beyond reward-prediction error. The RPEH is not the only available hypothesis about what type of information is encoded by dopaminergic activity in the midbrain (cf., Berridge, 2007; Friston et al., 2012; Graybiel, 2008; Wise, 2004). Current evidence does not speak univocally in favour of this hypothesis, and disagreement remains about to what extent the RPEH is supported by available evidence (Dayan & Niv, 2008; O'Doherty, 2012; Redgrave & Gurney, 2006). On the one hand, it has been claimed that "to date no alternative has mustered as convincing and multidirectional experimental support as the prediction-error theory of dopamine" (Niv & Montague, 2009, p. 342; see also Niv, 2009; Glimcher, 2011); on the other hand, the counter-claims have been put forward that the RPEH is an "elegant illusion" and that "[s]o far, incentive salience predictions [that is, predictions of an alternative hypothesis about dopamine] appear to best fit the data from situations that explicitly pit the dopamine hypotheses against each other" (Berridge, 2007, p. 424).

How has the RPEH become so successful then? What does it explain exactly? And, granted that it is at least intuitively uncontroversial that the RPEH is beautiful and elegant, in which sense can it be justifiably deemed deeper than alternatives? The present paper addresses these questions by firstly reconstructing the main historical events that led to the formulation and subsequent success of the RPEH (Section 2).

E-mail address: m.colombo@uvt.nl

1369-8486/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.shpsc.2013.10.006

With this historical account on the background, it is elucidated what and how the RPEH explains, contrasting it to the incentive salience hypothesis—arguably its most prominent current alternative. It is clarified that both hypotheses are concerned only with what type of information is encoded by dopaminergic activity. Specifically, the RPEH has the dual role of accurately describing the dynamic profile of phasic dopaminergic activity in the midbrain during reward-based learning and decision-making, and of explaining this profile by citing the representational role of dopaminergic phasic activity. If the RPEH is true, then a mechanism composed of midbrain dopaminergic neurons and their phasic activity carries out the task of learning what to do in the face of expected rewards, generating decisions accordingly (Section 3).

The paper finally explicates under which conditions some explanation of learning, motivation or decision-making phenomena based on the RPEH can be justifiably deemed deeper than some alternative explanation based on the incentive salience hypothesis. Two accounts of explanatory depth are considered. According to one account, deeper explanatory generalizations have wider scope (e.g., Hempel, 1959); according to the other, deeper explanatory generalizations show more degrees of invariance (e.g., Woodward & Hitchcock, 2003). It is argued that, although it is premature to maintain that explanations based on the RPEH are actually deeper—in either of these two senses of explanatory depth—than alternative explanations based on the incentive salience hypothesis, relevant available evidence indicates that they may well be (Section 4). The contribution of the paper to existing literature is summarised in the conclusion.

2. Reward-prediction error meets dopamine

Dopamine is a neurotransmitter in the brain.[1] It has significant effects on many aspects of cognition and behaviour, including motor control, learning, attention, motivation, decision-making and mood regulation. Dopamine is implicated in pathologies such as Parkinson's disease, schizophrenia, attention deficit hyperactivity disorder (ADHD) and addiction. These are some of the reasons for why so much work has been directed at understanding the type of information carried by neurons that utilize dopamine as a neurotransmitter as well as their functional roles in cognition and behaviour.

Neurons that use dopamine as a neurotransmitter to communicate information are called dopamine or dopaminergic neurons. Such neurons are phylogenetically old, and found in all mammals, birds, reptiles and insects. Dopaminergic neurons are localized in several brain networks in the diencephalon (a.k.a. interbrain), mesencephalon (a.k.a. midbrain) and olfactory bulb (Björklund & Dunnett, 2007). Approximately 90% of the total number of dopaminergic neurons is in the ventral part of the midbrain, which comprises different dopaminergic networks with separate pathways. One of these pathways is the nigrostriatal pathway. It links the substantia nigra, a structure in the midbrain, with the striatum, which is the largest nucleus of the basal ganglia in the forebrain and has two components: the putamen and the caudate nucleus. Another pathway is the mesolimbic, which links the ventral tegmental area in the midbrain to structures in the forebrain, external to the basal ganglia, such as the amygdala and the medial prefrontal cortex.

Dopamine neurons show two main patterns of firing activity, which modulates the level of extracellular dopamine: tonic and phasic activity (Grace, 1991). Tonic activity consists of regular firing patterns of ~1–6 Hz that maintain a slowly-changing, extracellular, base-level of extracellular dopamine in afferent brain structures. Phasic activity consists of a sudden change in the firing rate of dopamine neurons, which can increase up to ~20 Hz, causing a transient increase of extracellular dopamine concentrations.

The discovery that neurons can communicate by releasing chemicals was due to the German-born pharmacologist Otto Loewi—Nobel Prize winner in Physiology and Medicine along with co-recipient Sir Henry Dale—in 1921 (cf., Loewi, 1936). The discovery of dopamine as a neurotransmitter in the brain dates 1957, and was due to the Swedish pharmacologist Arvid Carlsson—Nobel Prize in Physiology and Medicine in 2000 along with co-recipients Eric Kandel and Paul Greengard (cf., Carlsson, 2003). Carlsson's work in the 1950s and 1960s paved the way to the findings that the basal ganglia contain the highest dopamine concentrations, that dopamine depletion is likely to impair motor function and that patients with Parkinson's disease have markedly reduced concentrations of dopamine in the caudate and putamen (cf. Carlsson, 1959, 1966).

Since at least the 1950s, the search for the mechanisms of reward-based learning and motivation has been taking place. James Olds and Peter Milner set out to investigate how electrical stimulation of certain brain areas could reinforce behaviour. They implanted electrodes in different areas of rats' brains and allowed them to move about a Skinner box. Rats received stimulation whenever they pressed a lever in the box. When this stimulation was targeted at the ventral tegmental area and basal forebrain, the rats showed signs of positive reinforcement, as they would repeatedly press the lever up to 2000 times per hour. These results suggested to Olds and Milner that they had "perhaps located a system within the brain whose peculiar function is to produce a rewarding effect on behavior" (Olds & Milner, 1954, p. 426).

The notion of "reward" here is to be understood within Thorndike's (1911) and Skinner's (1938) theories of learning. As Olds and Milner put it: "In its reinforcing capacity, a stimulus increases, decreases, or leaves unchanged the frequency of preceding responses, and accordingly it is called a reward, a punishment, or a neutral stimulus" (Olds & Milner, 1954, p. 419). So, some brain stimulation or some environmental stimulus is "rewarding" if animals learn to perform actions that are reliably followed by that stimulation or stimulus.

Later experiments confirmed that electrical self-stimulation of specific brain regions has the same impact on motivation as other natural rewards, like food or water for hungry or thirsty animals (Crow 1972; Trowill, Panksepp, & Gandelman, 1969). The idea that some neurotransmitter could be a relevant causal component of some mechanism of reward-based learning and motivation was substantiated by pharmacological studies (Stein, 1968, 1969). Based on subsequent pharmacological (Fibiger, 1978) and anatomical research (Lindvall & Björklund, 1974), hypotheses about the involvement of dopaminergic neurons in this mechanism began to be formulated. In Roy Wise's (1978) words: "[from the available evidence] it can be concluded that dopamine plays a specialized role in reward processes…It does seem to be the case that a dopaminergic system forms a critical link in the neural circuitry which confers rewarding qualities on intracranial stimulation…and on intravenous stimulant injections" (Wise, 1978, pp. 237–238).

Wise (1982) put forward one of the first hypotheses about dopamine function in cognition and behaviour that aimed to

[1] Neurotransmitters are chemicals that carry information from one neuron to another across synapses. Synapses are structures connecting neurons that allow one nerve cell to pass an electric or chemical signal to one or more cells. Synapses consist of a presynaptic nerve ending, which can contain neurotransmitters, a postsynaptic nerve ending, which can contain receptor sites for neurotransmitters, and the synaptic cleft, which is a physical gap between the presynaptic and the postsynaptic ending. After neurotransmitters are released by a presynaptic ending, they diffuse across the synaptic cleft and then bind with receptors on the postsynaptic ending, which alters the state of the postsynaptic neuron.