Dopamine, Prediction Error and Beyond

Kelly M. J. Diederen[1] and Paul C. Fletcher[2,3,4]

Abstract

A large body of work has linked dopaminergic signaling to learning and reward processing. It stresses the role of dopamine in reward prediction error signaling, a key neural signal that allows us to learn from past experiences, and that facilitates optimal choice behavior. Latterly, it has become clear that dopamine does not merely code prediction error size but also signals the difference between the expected value of rewards, and the value of rewards actually received, which is obtained through the integration of reward attributes such as the type, amount, probability and delay. More recent work has posited a role of dopamine in learning beyond rewards. These theories suggest that dopamine codes absolute or unsigned prediction errors, playing a key role in how the brain models associative regularities within its environment, while incorporating critical information about the reliability of those regularities. Work is emerging supporting this perspective and, it has inspired theoretical models of how certain forms of mental pathology may emerge in relation to dopamine function. Such pathology is frequently related to disturbed inferences leading to altered internal models of the environment. Thus, it is critical to understand the role of dopamine in error-related learning and inference.

Introduction

Dopamine is a critical modulatory neurotransmitter. Acting within distinct pathways, it is involved in a wide range of functions, including the control of movement, motivation, reward processing, and learning. Its perturbation has been linked to profound neurodegenerative and psychiatric impairments (Bissonette and Roesch 2016). In the past, dopamine's association with "happiness" or "pleasure" has been emphasized in view of its role in the prediction, anticipation, and approach behavior toward rewarding outcomes (Arias-Carrión and Pöppel 2007). Indeed, there is overwhelming evidence that dopamine guides learning about reward outcomes, by keeping track of violations in our expectations, called prediction errors (PEs) (Schultz 2016a). However, dopamine may have a role in the signaling of PEs that are not directly related to reward (Friston 2010). Here, we provide an overview of the evidence relating dopaminergic function to reward learning and discuss emerging work that suggests a crucial role for dopamine in predicting any future outcomes. In doing so, we consider how it may be a key contributor to setting up a model of associative regularities in the environment as a basis for flexible inference and how disruption in this role may parsimoniously explain key symptoms of neuropsychiatric disorders (Fletcher and Frith 2009; Sterzer and others 2018).

Prediction Errors

Predicting the outcomes of our actions is crucial for effective decision making and behavior. An efficient mechanism for learning is to keep track of violations in our expectations, termed PEs (outcome expected – outcome received) (Schultz 2016a). These errors effectively allow us to predict which outcomes are likely to be available at a particular time, and to guide our choices toward optimal behaviors. PEs are underpinned by dedicated neural signals which drive learning about outcomes in the domains of perception, motor function, punishment and reward (Den Ouden and others 2012). Reward PEs (RPEs) differ from sensory and motor PEs in that as well as engendering surprise (referred to as *unsigned* PEs), they indicate whether

[1]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[2]Department of Psychiatry, University of Cambridge, Cambridge, UK
[3]Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK
[4]Wellcome Trust MRC Institute of Metabolic Science, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

Corresponding Author:
Kelly M. J. Diederen, Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London SE5 8AF, UK.
Email: kelly.diederen@kcl.ac.uk
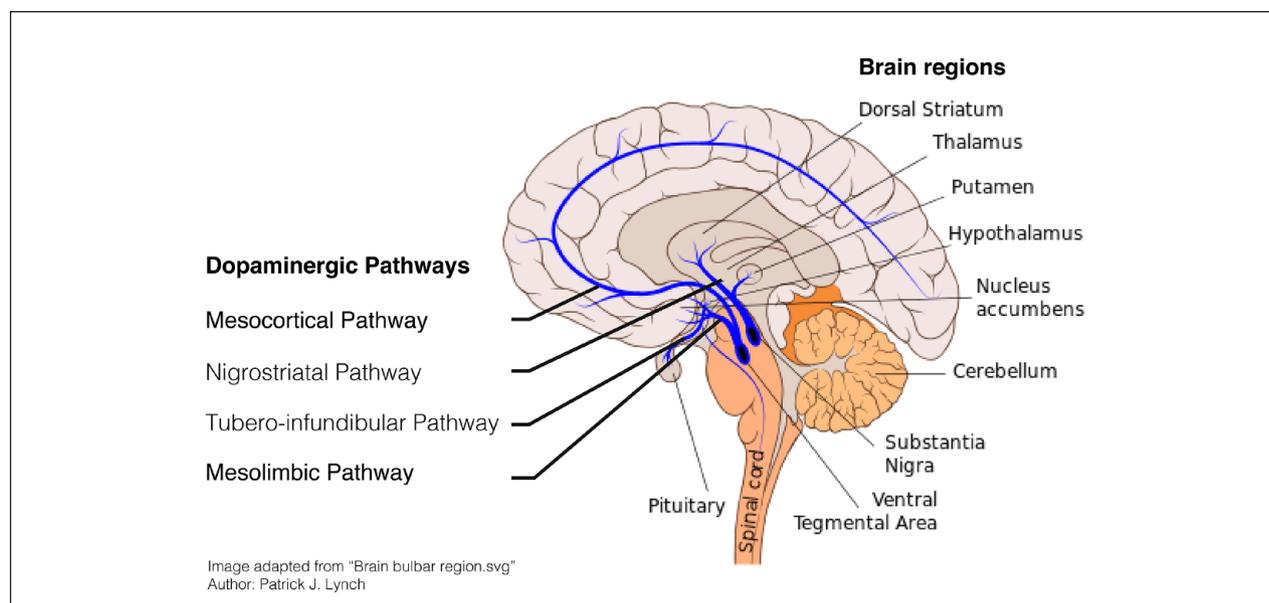
**Figure 1.** The dopaminergic system in the human brain.

outcomes were better or worse than expected, resulting in positively and negatively *signed* PEs (Den Ouden and others 2012). Signed and unsigned PEs are (at least partially) underpinned by separate neural substrates (Fiorillo 2013). While the neurotransmitter dopamine has consistently been shown to play a major in the encoding of RPEs (Schultz 2016a), its role in other domains is less clear.

Rewards. Neuroscience research broadly defines rewards as any positive, or pleasurable, outcomes, that we are motivated to obtain and that we will work for (Schultz 2016a). To determine the sign (positive/negative), and size of RPEs we need to know how much individuals value specific rewards (Levy and Glimcher 2012). Primary rewards, including food, drink, and sex, are innately valuable due to their intrinsic survival properties. By contrast, indirect rewards such as money, derive their positive value from their association (conditioned reinforcement) with pleasurable outcomes (Wise 2002). Primary and secondary rewards are associated with similar behaviors (e.g., choices) and dopamine responses, compatible with the idea that the brain transforms all rewards onto a single scale of value that facilitates decision making when different actions may procure different types of rewards (Lak and others 2014).

The Dopaminergic System. Dopamine is synthesized by dopamine neurons and thence transported via axonal projections widely throughout the brain. Although dopamine has multiple functions, the brain contains relatively few dopamine neurons; ~400,000 in the human brain which accounts for ~1% of the total neuronal population (Arias-Carrión and Pöppel 2007). The majority

of dopamine neurons are located in two small nuclei in the midbrain called the ventral tegmental area (VTA), and the substantia nigra (SN), which has two subnuclei called the pars compacta (SNc) and pars reticulata (SNr; Nair-Roberts and others 2008). The latter two so-called because of the presence of the dark pigment melanin within the dopamine neurons (Halliday and Törk 1986). Four functionally distinct dopamine projections can be distinguished (Fig. 1). RPE signaling is primarily facilitated by the mesolimbic pathway, which transmits dopamine from the VTA to the nucleus accumbens (NA) in the ventral striatum. By contrast, the nigrostriatal pathway, which connects the SNc to the dorsolateral striatum, and the premotor/motor cortex is thought to facilitate action-selection of the most rewarding action (García-García and others 2017).

Measuring Dopaminergic Function. In animals, the phasic (fast spiking) and tonic (slower responses) of VTA/SN dopamine neurons can be measured directly using single cell electrophysiology (Schultz and others 1997). These phasic dopamine responses lead to dopamine release in the NA, which can be measured at a lower temporal resolution using fast-scan cyclic voltammetry and microdialysis (Clark and others 2009; Hart and others 2014). Indirect measurements in animals consist of electrical stimulation of dopamine neurons and administration of drugs that act on the dopaminergic system (Olds and Milner 1954). More recently, optogenetic stimulation has been used in rodents and monkeys to directly stimulate dopamine neurons (Kim and others 2012; Stauffer and others 2016).
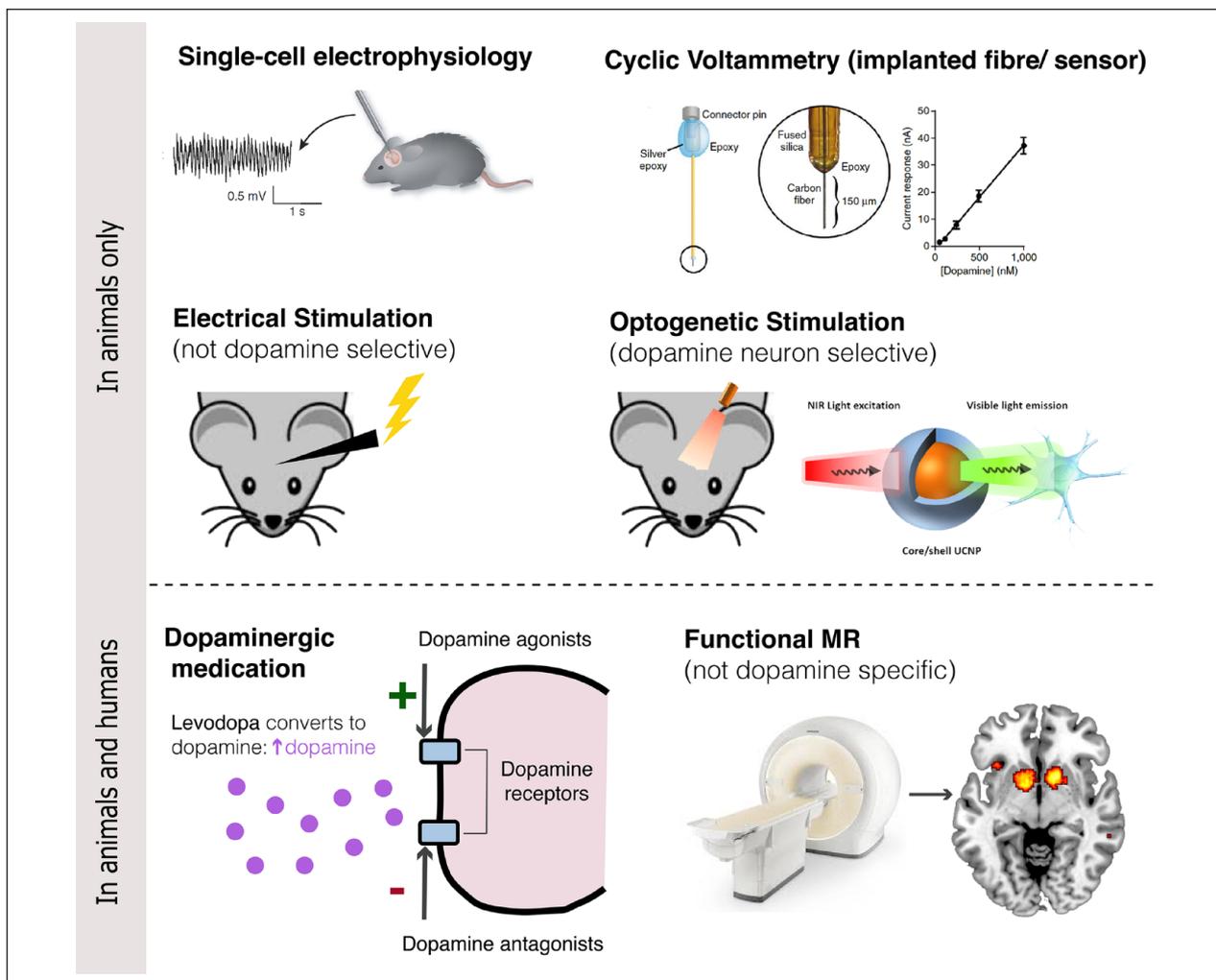
Figure 2. Techniques for investigating dopamine and prediction errors.

In humans, brain responses in the VTA/SN are mainly investigated using functional neuroimaging techniques (sometimes in conjunction with pharmacological challenges) such as functional MRI (fMRI) and position emission tomography (PET; Düzel and others 2015). fMRI measures changes in blood oxygen level, a proxy for neural activity, on a timescale of seconds. The low temporal and spatial resolution, however, make it difficult to determine whether observed signals reflect dopaminergic signaling as the VTA/SN are only partly made up of dopamine neurons, and the exact location of these nuclei varies across individuals (Düzel and others 2015). Although PET allows for non-invasive measurement of dopaminergic activity in humans, its temporal resolution is insufficient to draw direct comparisons to animal electrophysiological studies (Heiss 2009). In this review, we will discuss and integrate findings obtained using each of these techniques (Fig. 2), while considering the challenges in doing so.

RPEs and Reinforcement Learning. The idea of RPEs has long been central to ideas of classical and instrumental conditioning and to reinforcement learning (RL) generally. RL has its roots in the seminal work of Ivan Pavlov on classical (Pavlovian) conditioning in dogs (Pavlov and Anrep 1982), as well as in machine learning. Pavlov used the term reinforcement to describe the strengthening of the association between a reward (unconditioned stimulus [US]) and the conditioned stimulus (CS; e.g., sounding a bell). Repeated pairings of the CS with the US allowed Pavlov's dogs to learn to predict the availability of a reward (food) when they heard a bell, as indicated by the CS—salivation on hearing the bell.

Whereas Pavlov focused on situations in which the outcome (food) followed the conditioned stimulus (bell) irrespective of any behavioral reaction, Edward Thorndike (Thorndike 1898) and Burrhus Frederic Skinner (Skinner 1938) studied what has come to be known as instrumental or operant conditioning, in which the animal's behavior

determines whether the unconditioned stimulus is presented (Skinner 1963). A now famous experiment included placing hungry cats in an enclosed container, which Skinner referred to as a puzzle box, from which they had to escape in order to reach food. The first time a cat was placed in this situation it escaped only after it made the right action (pressing a lever) by chance. The time it took to perform this action decreased each time it was returned to the box, suggesting that the cat was learning, or, specifically, that the useful action was being reinforced. While classical (or Pavlovian) and instrumental conditioning entail rather different experimental set-ups, they are tightly related. Stimuli associated with reward through classical conditioning come to motivate behavior generally, and can, more specifically, motivate particular

behaviors that have been learned to be associated with the reward that they predict a phenomenon referred to as Pavlovian to Instrumental transfer (de Wit and Dickinson 2009; Estes 1948).

With regard to the involvement of RPE in RL, one key observation was that, if the PE is absent, learning does not occur even when a cue is strongly associated with an outcome. This is famously demonstrated in Kamin's blocking effect (Kamin 1969), in which a previously learned cue-outcome association (A → X) blocks the acquisition of learning when a new cue is added (AB → X). In this case, there is no PE to AB-X because A already predicts X and so, though B is associated with X, the association is not reinforced. This remarkable observation underpins formal RL models (see Box 1).

Box 1.

Reinforcement Leaning (RL).

RL. Formal RL models foster a more mechanistic understanding of the different computations that a neural system must solve to translate to changes in behavior. A first formal (computational) model of RL was developed by Robert Rescorla and Allan Wagner, termed the Rescorla-Wagner (RW) model (Rescorla and Wagner 1972), which specifies that learning slows as the reinforcer (reward) becomes more predicted as a function of decreasing RPEs:

$$y_n = y_{n-1} + \alpha * \delta_n \qquad (1)$$

Here, predictions of reward value (y) are updated iteratively (n) as a function of the size of the RPE (δ) and a constant, termed the learning rate (α) that determines the weight attributed to PEs to drive learning. if we expect to get £10, but receive £16 instead, we will have a PE of £6. If we have a learning rate of 0.5, we will update our next prediction of reward to equal £13 (£10 + 0.5 * 6). When the PE equals zero no more learning occurs (Niv and Schoenbaum 2008; Schultz and Dickinson 2000).
The RW model lacks a consideration of time, assuming that predictions of reward are specific to each individual trial and that trials are discrete. Noting this, Sutton and Barto (1998) introduced the idea that predictions are based on all future expected rewards (within a particular environment), with the additional feature that temporally closer rewards have more value than those in the more distant future. This led to the so-called temporal difference model (Sutton and Barto 1998), which is a form of dynamic programming, to include a discount factor (γ) in their calculation of the PE which determines the extent to which rewards that arrive earlier are more important than rewards that arrive later on:

$$\delta_n = r_n + \gamma \hat{V}_{n+1} 1 - \hat{V}_n \qquad (2)$$

where the PE (δ) on a specific trial indicate the difference between the expected value of all future reward ($\hat{V}$).
Another strategy for weighting earlier outcomes, and hence PEs, more than later ones, is to reduce the learning rate as trials progress. Decreasing your rate of learning as time progress is sensible as one's predictions become more reliable (and hence new outcomes less informative) as time progresses. Pearce and Hall (1980) allowed the learning rate to decrease across trials (α_n) as a function of the absolute PE ($|\delta|$) - which signals the extent to which previous predictions were wrong - and the learning rate on the previous trial, and an individually determined discount factor (γ):

$$\alpha_n = \gamma |\delta_{n-1}| + (1-\gamma)\alpha_{n-1} \qquad (3)$$

where the recursive process is initialized with the initial learning rate $\alpha_0 = \alpha$.
Extending RL Models to Decision Making. The above models speak directly to the relationship between cues and (predicted) outcomes. Models based on similar principles have extended the ideas to instrumental conditioning, where the context-dependent value of different action options must be tracked and used to optimize choice behavior. One simple but powerful instance of this is the Q-learning model (Dickinson and Balleine 1994; Sutton and Barto 1998). This is closely related to RL models based on classical condition as the relationship between choices and reward values is learned via the PE (δ). For each pair of stimuli, *A* and *B*, the model estimates the expected values of choosing *A*(Qa) and choosing *B*(Qb), on the basis of individual sequences of choices and outcomes. This value, termed a *Q* value, is essentially the expected reward obtained by taking that particular action. After every trial $n > 0$ the value of the chosen stimulus is updated according to the following rule:

$$QA_{n+1} = QA_n + \alpha * \delta_n \qquad (4)$$

and the PE is calculated using the following formula to indicate the difference between the predicted value of the chosen option and the maximum discounted future Q value:
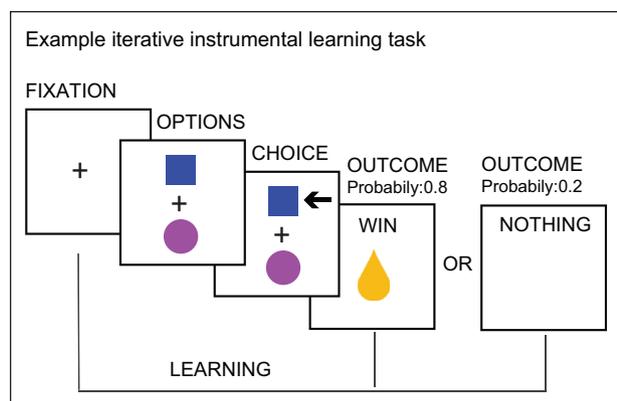
$$\delta_n = r_n - Q_n \tag{5}$$

Given the Q values, the associated probability of selecting each action is then estimated by implementing a softmax rule (see Sutton and Barto, 1998 for examples). The softmax rule has two parameters one denoting the learning rate and the second denoting the (inverse) temperature. The temperature specifies the noise or randomness in choice behavior.

Importantly, in all these models, learning only occurs if PEs are valenced, and they thus do not allow for learning associations of complex associative structures in the presence of non-RPEs.

In addition to the "model-free" RL models described above, there is a separate set of more flexible "model-based" RL models, which state that individuals build a cognitive model of environmental contingencies to allow for forward planning to identify the most rewarding options (Dickinson and Balleine 2002). Here, individuals evaluate possible actions by searching a cognitive model that represents the current *state* of the environment (e.g., the door is open), the likelihood that a reward will occur in this state, and how a decision may change the state (e.g., the door will close). Optimal decision making therefore requires individuals to predict future states, which can be learned from state PEs (Gläscher and others 2010).

Determining the Value of Rewards. Ultimately, the (model-free) RL models relate RPEs and learning to the value we assign to a particular reward (Sutton and Barto 1998). The question of how and why value is assigned is enormously complex. To identify the value of an expected reward individuals must integrate information on different reward attributes including its type, magnitude, probability, and timing (Lak and others 2014; Padoa-Schioppa 2011). Whereas reward value typically increases as the magnitude, probability, and temporal proximity increase, the weighting of each of these reward attributes varies across time and individuals. For instance, hunger increases the value of even a small or bland food reward. In addition, reward preferences vary across individuals, and depend on personality characteristics such as attention, motivation, patience and willingness to take a gamble (Padoa-Schioppa 2011). In general, people are risk avoidant, that is, they prefer smaller "safe" rewards over a gamble that can result in a larger, risky reward, but preference varies both across individuals and conditions (e.g., we are more likely to take a gamble when the amount of money at stake is small). Similarly, monkeys become more risk aversive (for juice rewards) when they are thirsty (Yamada and others 2013). In addition, our preference also depends on our ability to accurately learn about each of these reward characteristics and it has been shown that subjects' estimate of probabilities tends to be distorted (Stauffer and others 2015; Tobler and others 2008; Tversky and Kahneman 1974).

These insights have led researchers to investigate rewards in terms of subjective rather than objective values (Bartra and others 2013; Kahneman and Tversky 2013). Subjective values can be determined by an individual's choice behavior when asked to make a set of iterative choices between different options to determine the relative value of different rewards (Luce 2012; Taylor and Creelman 1967). The probability of choosing one option over others denotes the predicted value that subjects have attached to the available options. Crucially, choices are based on predictions of outcomes, which can be obtained through learning as the result of a choice is frequently not explicitly available to individuals (Cartoni and others 2013). In humans, we can additionally ask how much they prefer each reward or let them "play" in so-called first and second prize auctions in which people indicate how much of an endowment they are willing to pay for a particular reward (Becker and others 1964).

Dopamine and RPE Coding. In 1997, a clear neuronal substrate of RPEs was observed (Schultz and others 1997). Schultz and colleagues showed that dopamine neurons in the midbrain VTA/SN changed their firing rate in response to rewards, and to (conditioned) cues that are predictive of rewards. Specifically, dopamine neurons in the VTA/SN of macaque monkeys increased their firing rates for unexpected, but not expected, juice rewards. When a reward was preceded by a predictive visual stimulus, firing occurred in response to the stimulus but not the reward. This can be explained in terms of the predictive stimulus signaling a positive RPE and suggests that the visual cue had come to acquire properties of the reward itself. Moreover, when expected rewards were omitted the dopamine neurons showed a reduction in firing rates (indicating a negative PE). Overall, the findings clearly demonstrated that dopaminergic neurons did not respond to receiving a reward per se, but rather that these neurons tracked the violation in expected rewards. The authors furthermore showed that the observed dopamine responses obeyed the rules of RL models which provided further evidence for their key role in error-dependent learning. Numerous studies since then, including the measurement of dopamine release in the NA in animals,

Example iterative instrumental learning task

FIXATION

OPTIONS

CHOICE

OUTCOME
Probabily:0.8

OUTCOME
Probabily:0.2

WIN

OR

NOTHING

LEARNING

Figure 3. Example experiment for investigating reward prediction errors.

have proved compatible with these observations (Bayer and Glimcher 2005; Hart and others 2014). Importantly, it has recently been established that dopamine neurons do not inherit the RPEs from upstream regions but are directly involved in the computation of these PEs (Watabe-Uchida and others 2017).

In humans, the hypothesized role of dopamine in RPE signaling has been strongly supported by studies using neuroimaging and pharmacological dopaminergic manipulations. Using high-resolution fMRI, D'Ardenne and others (2008) studied the small VTA/SN nuclei in the human midbrain, as well as the NA. Activation occurred in response to unexpectedly large or small monetary rewards but not to rewards that were fully expected. Although these results do not necessitate that the observed PE signals were dopamine-dependent, complementary studies have shown that human PE signaling is modulated by administering (single) doses of dopaminergic agents to healthy individuals. Pessiglione and others (2006) found that the magnitude of RPE signals increased in the ventral striatum/NA of individuals who received levodopa (L-DOPA; a metabolic precursor of dopamine thought to increase dopamine signaling), compared to individuals who received haloperidol (a dopamine antagonist). Participants who received L-DOPA also won more money on the task, as they had learnt to choose the most rewarding option more frequently, suggesting that the elevated dopamine-dependent RPE drove improvements in learning, which optimized participants' decisions.

A large number of studies have confirmed the role of the VTA/SN and the striatum in the signaling of RPEs across both direct and indirect rewards (Garrison and others 2013; Sescousse and others 2013). See Figure 3 for an example of an RPE experiment in humans. Importantly, in humans the striatum was the key brain area encoding PEs (Fig. 4) in both instrumental and classical conditioning/reinforcement (Garrison and others 2013). Studies in

humans and animals alike have further established that the dopaminergic system integrates information about different reward characteristic, including the expected type (e.g., food or money), magnitude, probability, and time of reward (Fig. 5) to calculate RPEs (Diedederen and others 2016; Lak and others 2014; O'Doherty and others 2003; Tobler and others 2005).
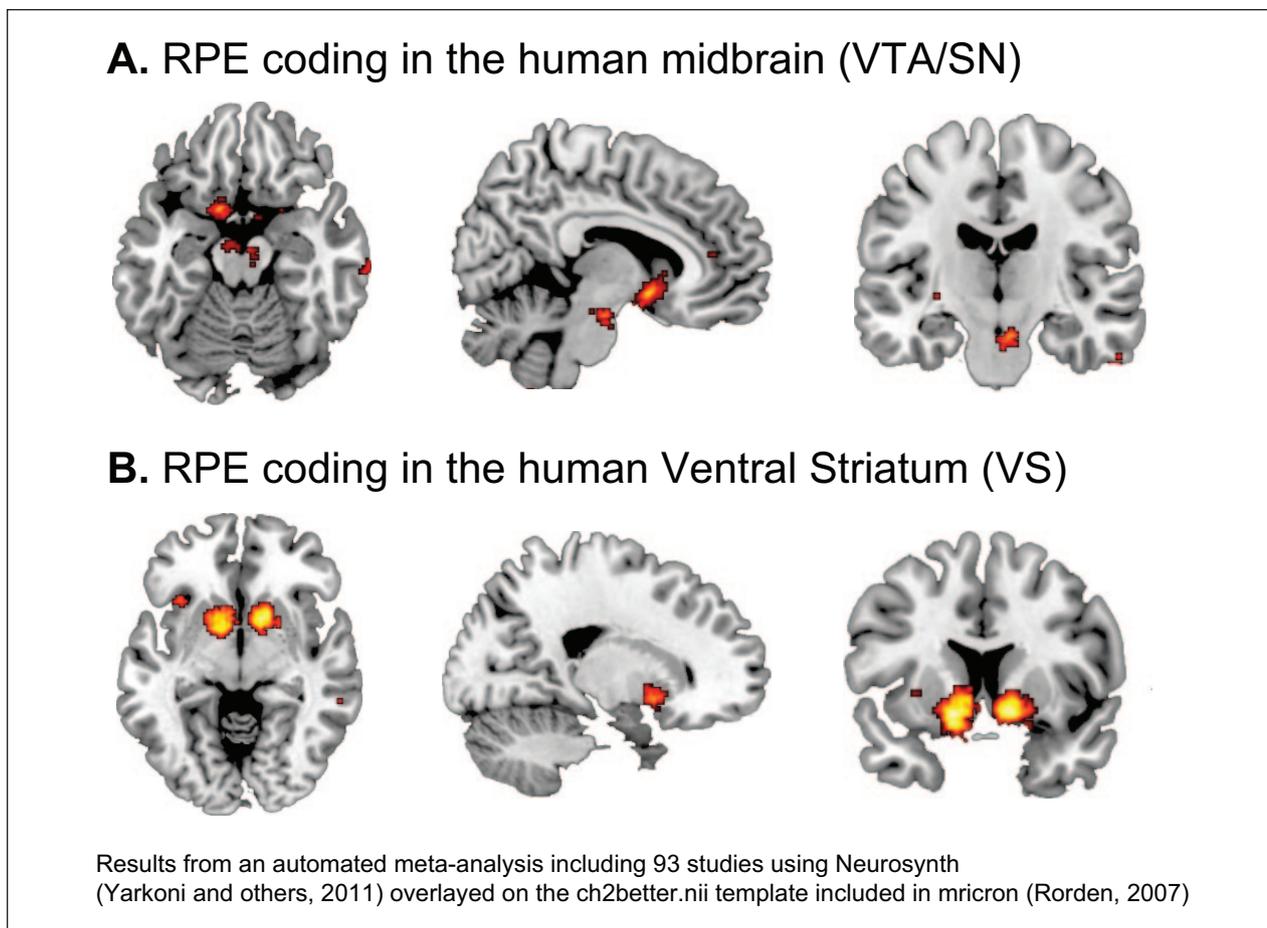
In addition, recent work in mice revealed that dopamine RPEs are sensitive to beliefs about the (model-based) "state" that an animal is in (Starkweather and others 2017). Each trial consisted of (1) a cue-reward state where the time until reward was drawn from a Gaussian distribution and (2) an interstimulus interval (ISI) state. In one task, odor cues predicted reward in 90% of trials, which meant that the transition from the cue-reward state to the ISI state was unobservable or *hidden*, and that longer cue-reward delays increased the belief that reward was omitted and that a state transition had occurred. Optogenetically identified dopamine neurons showed the highest responses to the latest rewards, suggesting that animals had inferred a state transition (to the ISI state) and no longer expected reward. The authors also found that a revised TD model that included a belief state, which tracks the probability of being in each state produced PE signals that resembled dopamine RPEs. In line with this, others found that administration of L-DOPA to healthy individuals increased model-based over model-free choice (Wunderlich and others 2012). However, other work did not find model-based state PEs in midbrain or striatal dopaminergic regions (Gläscher and others 2010).

It is important to note here that the encoding of RPEs is not limited to the VTA/SN as other regions show responses depending on the nature for reward (Garrison and others 2013; Sescousse and others 2013). Specifically, whereas monetary RPEs were additionally observed in the orbitofrontal cortex, food and erotic rewards additionally engaged the anterior insula and the amygdala (Sescousse and others 2013).

Dopamine Beyond RPE Coding

Salience

In addition to its role in RPE coding, there is incomplete evidence that dopamine encodes salience, that is, the extent that a stimulus is particularly noticeable (Schultz 2016b). Focusing attention on stimuli that stand out could be evolutionary advantageous, as it directs attention to those stimuli that are likely to be of importance, for example, noticing a potential predator. It is important to note that RPE coding and salience are not mutually exclusive as the experience of any type of PE, including RPEs, is salient. As such, salience accounts propose a broader role for dopamine than RPE coding. Different types of salience

A. RPE coding in the human midbrain (VTA/SN)

B. RPE coding in the human Ventral Striatum (VS)

Results from an automated meta-analysis including 93 studies using Neurosynth
(Yarkoni and others, 2011) overlayed on the ch2better.nii template included in mricron (Rorden, 2007)

Figure 4. Reward prediction error (RPE) coding in the human brain.

have been defined and we consider these below, taking the view that the different forms of salience may relate to the extent to which a stimulus has been processed. In addition, we discuss the proposed role of dopamine in signaling specific salient events, including identity PEs and novelty.

Physical Salience

The debate on the role of dopamine in attributing salience focuses in part on whether salience attribution is limited to events that are likely associated with rewards. One line of evidence shows that physically salient sensory stimuli, such as tones and lights, evoke very rapid (50-110 ms), phasic excitations in dopamine neurons (Comoli and others 2003; Dommett and others 2005). This rapid response does not allow detailed identification and evaluation of the stimulus and is therefore unlikely to provide information about a potentially associated reward, although salient and novel stimuli might become erroneously associated with reward (Fiorillo 2013). Novel and physically salient stimuli might, however, be inherently rewarding as they provide unexpected, new information, that might

be of value for adaptive behavior (e.g., noticing a brightly colored object in a tree that might indicate an appetitive food; Daw and others 2002b; Reed and others 1996).

Novelty

As introduced above, a particular type of salient stimuli that recruit dopaminergic responses relate to novelty (Rangel-Gomez and Meeter 2016). For instance, microdialysis studies showed that novel stimuli can evoke dopamine release (Bassareo and others 2002). Dopamine neurons increase their responses in the face of novelty; once novel stimuli become familiar and are not reinforced, dopamine responses habituate (Schultz, 1998). This raises the question whether novelty responses occur purely because of their salient properties. In line with this notion, pharmacological dopaminergic challenge can speed up, and enhance early novelty detection, but does not affect further processing of novel stimuli (see Rangel-Gomez and Meeter 2016, for a review). Bunzeck and Düzel (2006), however, found that whereas the fMRI signal in the SN/VTA responded to novel stimuli, no such
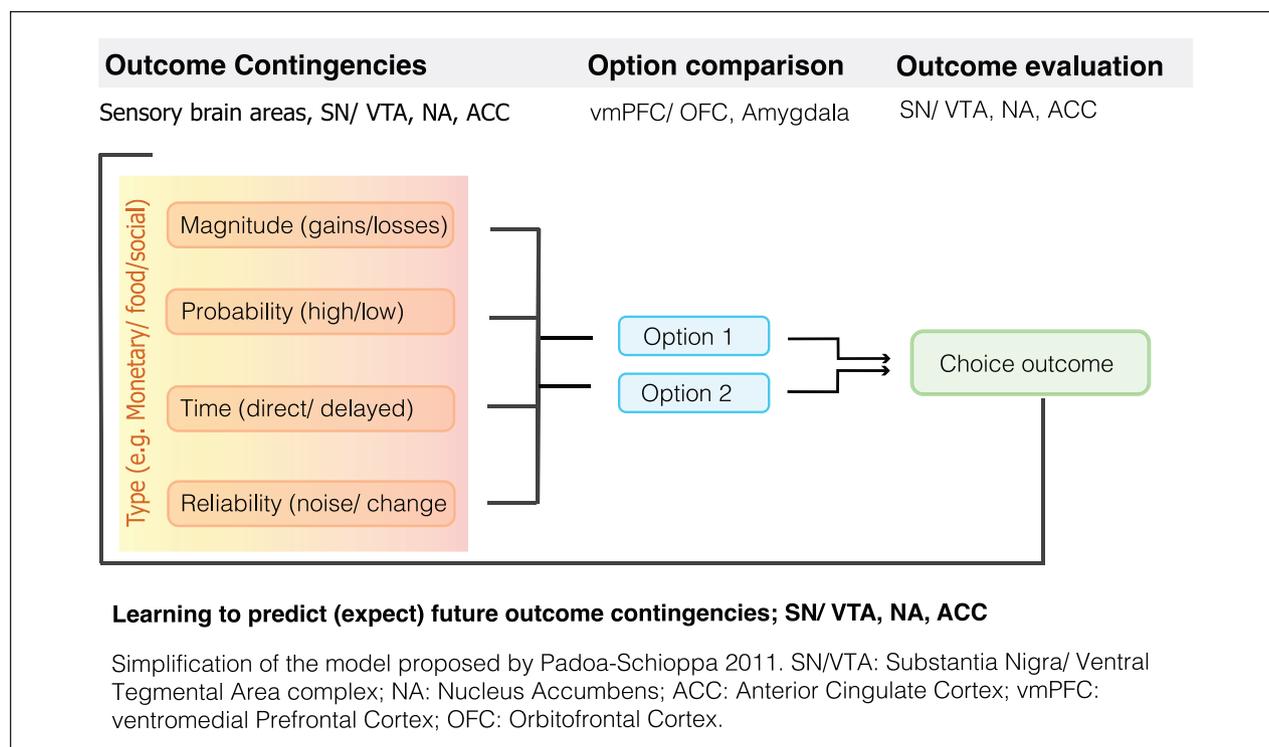
Figure 5. Reward learning and decision making.

effect could be found for other types of salience, including rareness and negative emotional valence, suggesting that dopamine might be particularly responsive to novelty. A later study found that SN/VTA responses to novel stimuli only occurred when novel stimuli were unexpected, bearing close resemblance to findings about reward, which show that responses to reward only occur when unexpected, thus signaling a PE. Although the exact role of dopamine in response to novelty is yet to be determined, it has been suggested, that novelty may motivate exploration which could result in higher rewards (Düzel and others 2010; Kakade and Dayan, 2002; Suri and others, 2001, but also see Lisman and Grace 2005).

Surprise Salience

Surprise salience, often called surprise, reflects the extent to which a more fully processed stimulus is unexpected (Ungless 2004). As such it operates at a cognitive rather than at a (purely) perceptual level. Surprise can, for instance, denote the magnitude of the PE, independent of its valence (positive/negative). It is thought that this surprise (or *unsigned*) PE signal indicates the degree to which an outcome is unexpected, independent of its sign, and thereby controls the rate of learning (Pearce and Hall 1980), whereas the signed RPE signals the extent to which an outcome is better or worse than expected (Rescorla and Wagner 1972; Sutton and Barto 1998).

A recent meta-analysis across human fMRI studies revealed support for a surprise-encoding network, including the anterior cingulate cortex (ACC), insula and dorsal striatum (Fouragnan and others 2018). Neurophysiological evidence suggests that unsigned PEs, are mainly coded in the cortex, including the dorsal ACC (Hayden and others 2011). This finding is confirmed in human studies that observed prefrontal PE coding in causal learning tasks, in the absence of explicit rewards (Corlett and others 2004; Fletcher and others 2001; Turner and others 2004). Although these brain regions receive dopamine projections (Esber and others 2012) these findings cannot allow inference about the role of dopamine in the encoding of surprise. To identify a potential role of dopamine in coding unsigned PEs (in this case, responses that were indistinguishable for positive and negative PEs), we used dopaminergic perturbations and showed that the dopamine antagonist sulpiride selectively decreased the encoding of unsigned PEs relative to reward reliability in the human superior frontal cortex (Haarsma and others 2019), but not in the striatum or midbrain, which was specific for RPEs (Diederen and others 2017; Haarsma and others 2019).

It should be noted though that, in this study, all PEs occurred in a reward context and it is not clear whether brain responses that do not distinguish negative and positive PEs might be different from responses to PEs that are unrelated to rewarding outcomes. It is conceivable

that the potential sensitivity of dopamine to unsigned PEs is still geared toward rewards and would not occur for unrewarded stimuli (Fiorillo 2013).

Identity and Sensory PEs

A further interesting observation comes from investigations of brain responses to rewards that are matched in (expected) value but differ in reward type/identity (e.g., receiving an equally valued pear instead of the expected apple). Such identity or sensory PEs engender surprise and can as such be considered salient. Using sensory preconditioning (associations between different neutral stimuli) and optogenetics in rodents, recent work showed that the acquisition of information about transitions between non-rewarding events is also driven by PEs and that, dopamine transients were sufficient to support this type of learning (Sharpe and others 2017). These findings were confirmed by recent work in both humans and rats which observed PE signals when the identity of the expected reward was violated (different odors), but the value was kept identical (Howard and Kahnt 2018; Takahashi and others 2017). Interestingly, in the work by Takahashi and others, dopamine responses to changes in value and identity did not occur in different neuronal populations. In contrast, recent work found distinct RPE and identity PE signals in the human midbrain (Boorman and others 2016). However, studies using cyclic voltammetry to monitor dopamine release failed to observe identity PEs (Collins and others 2016; Papageorgiou and others 2016). More work, including studies on fast phasic responses of dopamine neurons is needed to further investigate a potential for dopamine in signaling identity PEs, and to directly contrast work across different techniques and species.

Motivational Salience

Motivational salience refers to the quality that drives approach behavior for rewarding outcomes and avoidance behavior for aversive outcomes once the physical salience, surprise, and RPEs associated with a stimulus or option have been processed (Robinson and Berridge 2008). Such salience attribution would occur in between the identification of reward, and the generation of action to pursue it (McClure and others 2003). A role for dopamine in aversive salience is, however, heavily contested (Fiorillo 2013).

While neurons in the non-human primate SN increase their firing rate at very short latencies to unexpected stimuli, independent of whether they were rewards or punishment (Matsumoto and Hikosaka 2009), some have reinterpreted this finding as reflecting the physical intensity of stimuli, not their aversiveness (Fiorillo 2013).
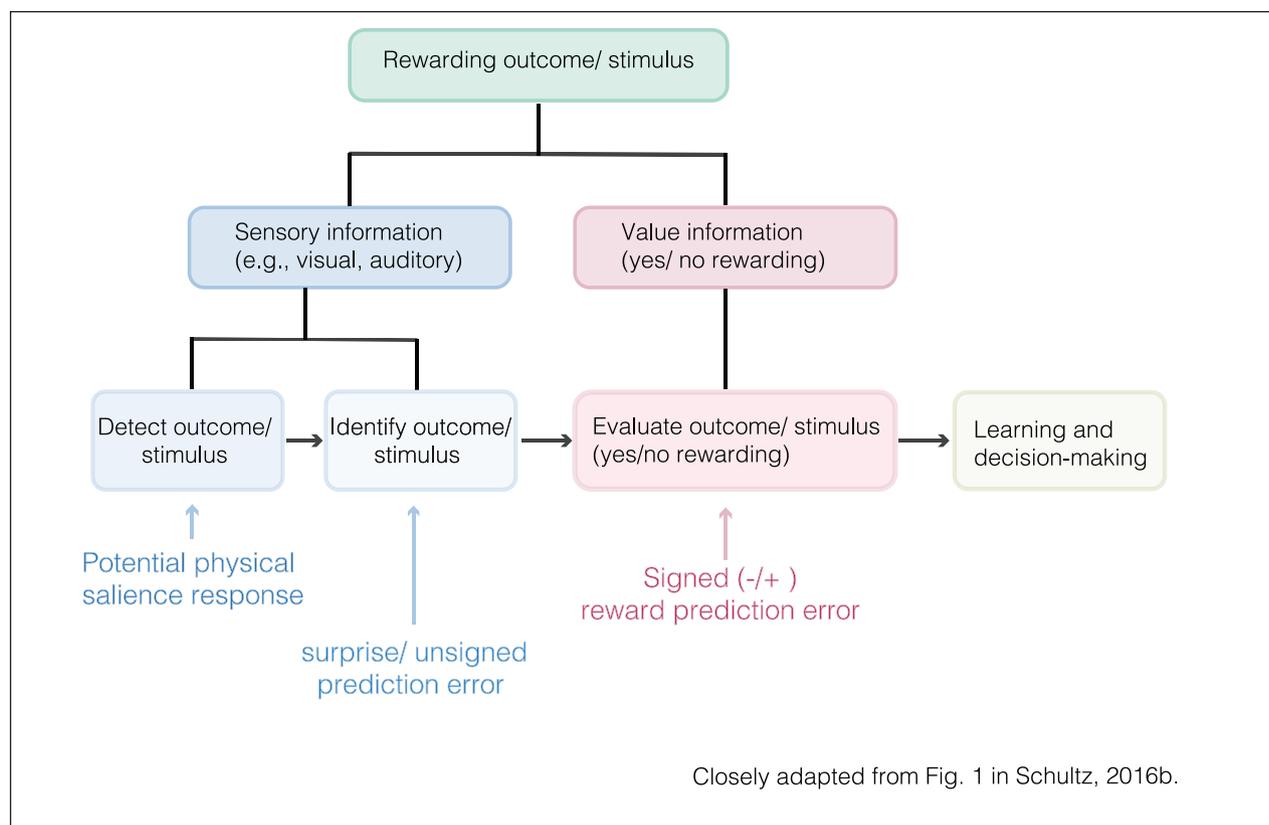
Others have found that aversive stimuli increase firing in a minority of midbrain neurons in the SN/VTA at longer latencies (Chiodo and others 1980; Mantz and others 1989). To shed more light on these findings, Ungless and others (2004) studied the properties of midbrain neurons that showed aversive responses and found that these midbrain neurons were not dopaminergic. In addition, the authors observed that neurochemically identified dopamine neurons decreased their firing to aversive stimuli. In addition, Fiorillo (2013) observed evidence that supported the existence of opponent neural representations for reward value and aversive outcomes (punishment), which the author concluded to be indicative of the existence of four types of value-sensitive neurons corresponding to reward-ON, reward-OFF, aversive-ON, and aversive-OFF of which only reward-ON was clearly dopamine-mediated (Fiorillo 2013). This is in line with earlier work that showed that motivationally salient events such as the unexpected omission of reward and the unexpected presentation of a stimulus predicting reward omission inhibit dopamine neurons (Tobler and others 2003). Finally, it has been argued that the relieving omission of an expected aversive stimulus can be considered a reward and might therefore evoke dopaminergic responses (Daw and others 2002a; Solomon and Corbit 1978). As such, it appears that dopamine responds selectively to (potentially) positively valenced outcomes, which is formulated in the notion of incentive salience (Robinson and Berridge 2008).

Robinson and Berridge (2008) suggested that mesolimbic dopamine is selectively involved in attributing incentive salience to potential objects or options to guide approach behavior, and that it has no role in RPE coding. Specifically, the authors argue that blocking dopamine selectively inhibits reward-seeking actions, without affecting valuation and the associated RPE of an outcome. This is in strong contrast to the overwhelming evidence for dopamine in RPE coding, and it has been argued by many that dopamine plays a dual role, which guides learning from RPEs and ongoing approach behavior (McClure and others 2003; Schultz 2016b).

Overall, there is relative consensus that mesolimbic dopamine plays a role in the attribution of physical and surprise salience (Schultz 2016b). It is, however, contested whether dopamine processes (motivationally) salient stimuli that are unlikely to be rewarding (i.e., aversive), and whether salience and PE accounts are mutually exclusive (Daw and others 2002b; Robinson and Berridge 2008).

Integrating RPEs and Salience Accounts of Dopamine

There are several accounts integrating the proposed roles of dopamine in reward prediction (error coding) and

```
                    ┌─────────────────────────────┐
                    │  Rewarding outcome/ stimulus │
                    └─────────────────────────────┘
          ┌───────────────────────┐   ┌──────────────────────┐
          │ Sensory information   │   │ Value information     │
          │ (e.g., visual,        │   │ (yes/ no rewarding)   │
          │  auditory)            │   │                       │
          └───────────────────────┘   └──────────────────────┘

  ┌──────────┐   ┌──────────┐   ┌──────────────────┐   ┌──────────┐
  │ Detect   │ → │ Identify │ → │ Evaluate outcome/│ → │ Learning │
  │ outcome/ │   │ outcome/ │   │ stimulus         │   │ and      │
  │ stimulus │   │ stimulus │   │ (yes/no rewarding)│  │ decision-│
  └──────────┘   └──────────┘   └──────────────────┘   │ making   │
                                                         └──────────┘
```

Potential physical salience response

surprise/ unsigned prediction error

Signed (-/+) reward prediction error

Closely adapted from Fig. 1 in Schultz, 2016b.

Figure 6. Sequel identification of rewarding outcomes/stimuli.

salience. For instance, Schultz (2016b) concluded that dopamine neurons have a "two component response" which integrates accounts of physical salience and RPE coding. Specifically, the first, rapid, component consists of a transient unselective response to a large variety of unexpected stimuli or events, whereas the later, less transient, response signals the occurrence of an RPE (Fig. 6).

In addition, formal learning models such as hybrid RW-PH RL models (Box 1) include a role for surprise as well as the RPE. Here, RPEs drive the trial-wise extent of learning, whereas surprise drives changes in the learning rate across time. Importantly, such hybrid models better predict individuals' learning behavior than these models alone (Diederen and Schultz 2015; Li and others 2011).

Furthermore, investigations in learning models that include choice behavior (Box 1), indicate that increases in dopamine activation resulting from increases in positive PEs increases the likelihood of choosing an action that leads to reward (Sutton and Barto 1998). Consistent with this, in addition to the phasic response of dopamine neurons in the SN/VTA, dopamine release in the striatum facilitates synaptic plasticity and can directly modulate reward-seeking behavior (Phillips and others 2003; Wickens and others 2003). Thus, it seems that accounts of physical salience, surprise, and incentive (but not

aversive) motivational salience, are compatible with RPE accounts as each of these processes appear to be integrated (Daw and others 2002b; Schultz 2016b).

Drawing on recent advances in artificial intelligence, Wang and others (2018) proposed a neurologically plausible meta-reinforcement account where dopamine-driven synaptic plasticity can train a more general and efficient learning system in the prefrontal cortex, allowing it to generalize its learning across different tasks and contexts. The authors carried out a set of simulations that provided support for this account, however, work involving experimental data is required to further test this model.

In addition, some authors have hypothesized that dopamine neurons serve a far more general function in signaling the expectation of (any) information (Bromberg-Martin and Hikosaka, 2009) or signaling errors in any type of prediction where value is only one of the dimensions (Langdon and others 2018; Takahashi and others 2017). The latter bears similarity to the notion that dopamine signals PEs independent of the domain in which these errors occur, thus allowing them to support a broader range of learning.

When combining findings, and theories of dopamine function to date, there appears to be little doubt that dopamine encodes RPEs (but see Friston and others 2015),

whereas the hypothesized role in salience coding is slightly more contested. The main question though is whether dopamine *uniquely* codes RPEs, or whether this is one of the (many) functions of dopamine. In light of the work discussed above, the most pressing enquiry is to establish whether dopamine neurons might compute any type of PE, independent of its domain, and how dopamine neurons interact with other brain systems to support other types of learning. The first question could be addressed by testing PEs in different domains, using tasks in which value is held constant or absent. In addition, it would be important to test for several types of PEs at the same time to scrutinize accounts of "multidimensional" PE signaling (Langdon and others 2018).

Dopamine PEs as a General Mechanism for Learning and Inference

Recent theories relating to the predictive processing framework and active inference have postulated an entirely different role for dopamine (Bastos and others 2012; Friston 2010; Friston and others 2012; Friston and others 2015). According to these theories, which are embedded in Bayesian (inference) models and based on early cybernetic theories, the brain is a predictive "machine" that updates its "model of the world" when PEs occur and its expectations are violated (Ashby 1952; Bayes and others 1763). Note that these accounts differ from Bayesian inference models as the latter characterize "the (computational) problem" that individuals are trying to solve, without making any explicit claims about their neurocognitive architecture (Griffiths and others 2012; Jacobs and Kruschke 2011). In contrast, predictive processing theories often additionally propose specific ways in which Bayesian inference may be implemented in the brain.

Although these ideas resemble RL accounts of learning, these novel theories differ in some major respects. First, they go beyond reward learning, seeing PEs as a generic model for inference and learning (termed belief updating) across sensory and cognitive domains. Furthermore, in these models behavior is optimal, not when reward is maximized, but rather when surprise (or the PE) is minimized. Moreover, they state that dopamine does not code an RPE, but rather codes the precision or reliability of the PE. These models also directly link perception and beliefs with both engaged in making sense of inputs by inferring their causes, an idea that goes back to Von Helmholtz (Von Helmholtz 1867). As PEs indicate unexpectedness in these models, they have clear links with salience accounts of dopamine.

In the predictive processing framework, there is a hierarchy where lower-level PEs signal violations of a sensory nature, while higher order PEs signal violations of

beliefs about the probabilistic structure of the environment and its volatility (inverse stability; Friston and others 2014). PEs emitted by a lower-level system becomes the input for a higher level system, whereas feedback from the higher-level system provides the prior beliefs for the lower level system.

The active inference account extends predictive coding into the domain of action and motor control (Friston and others 2012; Friston and others 2015). In simple terms, in this model, surprise can not only be minimized by improving one's "model of the world" but also by those actions that have predictable outcomes. More formally, perception and action can, respectively, minimize exteroceptive and proprioceptive PEs. This helps individuals to avoid exchanges with the environment that might be harmful (see Friston and others 2012 and Friston and others 2015 for details).

The precision weighting of PEs has been suggested to be dopamine dependent and to ensure that neural systems encoding predictions errors respond more strongly when new information is more reliable (i.e., minimizing surprise) and hence more informative. Lower precision can result from unclear predictions (e.g., early in learning), from noisy perceptual stimuli (e.g., conditions of poor visibility), high variability in the association between different outcomes, and changes in previously learnt associations or environmental volatility (Adams and others 2013; Bastos and others 2012; Friston 2009). Whereas previous work on reward learning has shown that PEs are coded relative to their precision (uncertainty), and that dopaminergic perturbations can modulate the precision-weighting process, a key distinction is that precision weighting is only included in some RL models (e.g. the Pearce-Hall model), whereas it is a key element of Bayesian inference models.

Although many studies have demonstrated support for approximate Bayesian inference, there is little direct experimental evidence to substantiate predictive coding, processing, and active inference frameworks. Some preliminary work using fMRI and PET found that belief updating activated the human VTA/SN and striatum independent of surprise or RPEs, and that these brain responses correlated with midbrain dopamine D2/D3 receptor availability and striatal dopamine release capacity (Nour and others 2018; Schwartenbeck and others 2016). This however contrasts with earlier work that specifically implicated the anterior cingulate cortex in belief-updating (O'Reilly 2013). Other work has more explicitly investigated the proposed hierarchical nature of PE coding, and revealed that whereas precision-weighted low-level PEs (stimulus associations) were coded by the VTA/SN, a higher level precision weighted PEs (expected changes in stimulus associations) engaged brain areas thought to be modulated by the neurotransmitter acetylcholine

(Diaconescu and others 2017; Iglesias and others 2013; Payzan-LeNestour and others 2013; Yu and Dayan 2005).

Although these studies provide initial support for a role of dopamine in coding PEs beyond reward, and playing a critical role in the building and updating of internal models of the world, it is important to establish the role of dopamine in this process as the SN/VTA contain dopaminergic and non-dopaminergic neurons (Nair-Roberts and others 2008). Furthermore, the findings that higher level PEs occurred in areas modulated by acetylcholine suggests that the dopamine might not be the primary neurotransmitter for coding PEs across the hierarchy.

PEs and Other Neuromodulary Systems

It is well known that dopamine interacts with other neurotransmitters, and that there is evidence of other neurotransmitters coding PEs. As a comprehensive overview of this topic is beyond the scope of this review, we illustrate this with the use of a few, select, examples.

Although much of the evidence is indirect, glutamate has frequently been associated with PE coding (see Pennartz and others 2000 and Lapish and others 2006, for theoretical accounts). In line with the notion that NMDA receptors drive dopamine responses to positive PEs, Jocham and others (2014) found that the NMDA antagonist memantine reduced positive but not negative PEs in the human striatum. In contrast, using the NMDA receptor antagonist ketamine, Corlett and others (2006) observed perturbed PE coding in the frontal cortex, but not in the striatum. Using a different approach, White and others (2015) found that glutamate in the SN, measured using MR spectroscopy, correlated with PE signals, in healthy individuals.

Noradrenaline and serotonin have also been implicated in PE signaling. Bouret and Sara (2004) found that noradrenergic neurons of the locus coeruleus in rats showed an RPE response similar to that observed for dopamine. Serotonergic neurons on the other hand, coded the magnitude of the PE (i.e., unsigned PEs) but did not differentiate between positive and negative PEs (Matias and others 2017).

Other neurotransmitters have been shown to interact with dopamine neurons to facilitate learning. For instance, GABA (γ-aminobutyric acid) neurons inhibit dopamine neurons when reward is expected, with contributes to the calculation of RPEs (Eshel and others 2015), a finding that was confirmed by Sharpe and others (2017). Furthermore, Kempadoo and others (2016) showed a tight relationship between dopamine and noradrenergic in facilitating learning. Finally, acetylcholine and norepinephrine have been proposed to signal uncertainty (Yu and Dayan 2005), which is a key component in learning and novel accounts of inference (see previous section).

Clinical Implications: The Case of Psychosis

A deeper understanding of the precise role of dopamine in PE coding and related functions is not merely of theoretical interest as dopamine dysfunction has been implicated in a range of diseases, including Parkinson's disease, Huntington's disease, substance use disorders, depression, anxiety disorder, and attention hyperactivity deficit disorder. There is particularly powerful evidence linking altered dopamine function to psychosis in the context of illnesses such as schizophrenia (Gatt and others 2015; Kollins and Adcock 2014; Nestler and Carlezon 2006). Although it is unclear how exactly altered dopamine can give rise to the symptoms of psychosis, several mechanisms have been put forward. Several proposed mechanisms have drawn on alterations in one or more of the dopaminergic mechanisms discussed in this review. As a broader discussion is out of the scope of this review, here we will provide a brief overview of dysfunctions in RPE coding and related concepts in psychosis, and some of the theories that have been put forward to link these dysfunctions to psychosis.

Psychosis has consistently been linked to increased presynaptic dopamine in the striatum (Howes and others 2012), with the dopaminergic alteration possibly preceding the onset of clinical-level psychosis (Howes and others 2011). In addition to these findings, people with psychosis present with an increased density in striatal dopamine D2 receptors, and alterations in genes involved in dopamine function (for a recent review, see McCutcheon and others 2019). Given this, and the fact that the primary treatment for psychosis is dopamine blockade, there has been a growing interest, beginning perhaps with the work of Robert Miller (1976) in embedding the understanding of the basic neuroscience of dopamine into models of psychosis. Indeed, multiple theorist have referred to dopamine as the final common pathway to psychosis (Howes and Murray 2014; Maia and Frank 2017), an account that was supported in a recent review (Valton and others 2017).

Psychosis has been associated with dysfunctions in a number of the dopamine-mediated processes described above (see Deserno and others 2013; Radua and others 2015; Maia and Frank 2017, for overviews). In brief, people with psychosis show attenuated behavioral and neural responses to reward predicting cues, whereas responses to neutral (or irrelevant) cues are increased, suggesting that these individuals experience difficulty identifying predictors of valuable outcomes. In the domain of RL, both patients on and off dopaminergic antipsychotic medication experience difficulties learning from positive RPEs, which is paralleled by attenuated coding of neural RPE signals, while learning from negative RPEs and their accompanying neural responses are preserved.

An influential attempt to explain how dopaminergic alterations may produce psychosis uses the notion of aberrant salience (Kapur 2003). Simply put, the idea is that an erratic phasic dopaminergic signal is experienced as an altered experience of the significance of environmental events and stimuli, which in turn drives a new appraisal of one's environment and the ensuing altered, and, apparently irrational and inexplicable beliefs (delusions). This "aberrant salience" framework has inspired subsequent cognitive and neuroimaging work supporting the view that irrelevant stimuli may be imbued erroneously with salience in people with psychosis (Murray and others 2008; Roiser and others 2009). However, there is much that remains unexplained by this initial theory and later models have sought to develop more comprehensive explanations of how dopamine function may drive shifted experiences of the world. In this regard, the important insights offered by studies of its role in PE-driven RL have proven very fruitful.

Recent mechanistic accounts have extended previous theories to explain how altered dopamine may give rise to both the positive (hallucinations and delusions) and negative symptoms of psychosis (e.g., affective flattening, alogia, and avolition; Deserno and others 2013; Maia and Frank, 2017). In brief, the authors proposed that reduced dopamine firing for relevant stimuli underlies negative symptoms, whereas an increase in spontaneous phasic dopamine release leads to excessive responses to neutral stimuli and PEs. Deserno and others (2013) furthermore hypothesized that aberrant PEs encode non-salient events as surprising, which drives aberrant learning resulting in these events being imbued erroneously with high incentive values, which can lead to positive psychotic symptoms. Negative symptoms on the other hand, are thought to result from a failure to use PEs to obtain accurate estimates of value. For a detailed mechanistic account at the computational and neurobiological level, see Maia and Frank (2017).

In recent years, it has been increasingly theorized that altered Bayesian inference could explain the symptoms of psychosis (see Valton and others 2017 and Heinz and others 2019, for an overview). Some of these accounts are (relatively) agnostic about the neural mechanisms underlying the hypothesized deficits in inference (Fletcher and Frith 2009; Valton and others 2017). Others have specified how altered Bayesian inference might be implemented in the brain by incorporating principles of predictive coding, predictive processing or active inference (Adams and others 2013; Jardri and Deneve, 2013). In multiple of these accounts, the idea is that the critical impact of dopaminergic perturbation lies in a change in the experienced precision of the PE signal, giving it undue weight and making it possible to change even long held and widely shared beliefs into the odd beliefs observed in psychosis. To date, support for these models has mainly been provided through simulations, stressing the need for more experimental work in individuals with psychosis (see Valton and others 2017 and Heinz and others 2019, for a summary of preliminary evidence).

It is important to remember that these "bayesian predictive coding" accounts entail the conjoining of two separate theoretical approaches. One (Bayesian inference) relates to a system's computational goal while the other, predictive coding, is an "algorithmic motif" (Aitchison and Lengyel 2017). A predictive coding system does not necessarily engage in Bayesian inference while Bayesian inference does not necessarily entail predictive coding (see Aitchison and Lengyel 2017 for a comprehensive discussion). Thus, many of the assumptions of the above approaches have yet to be empirically validated. Nonetheless, they do offer opportunities to credibly link emerging insights into dopaminergic contributions to PE to our attempts to understand the complex and baffling symptoms and subjective experiences of psychosis.

Conclusions

A wealth of studies has confirmed the role of dopamine in RPE coding, using a large range of different techniques and validation across different species, including rodents, monkeys, and non-human primates. It is, however, likely that dopamine has an additional role in signaling the amount of surprised associated with a rewarding outcome or stimulus, which can well be integrated with the RPE framework. Finally, novel models suggest an entirely different role for dopamine in PE coding across the whole brain, which might be exceptionally important for understanding clinical conditions associated with altered dopamine processing such as psychosis. It will be crucial to investigate these novel accounts in future studies using experimental designs and data, and to further the work on the link between these models and clinical conditions.

ORCID iD

Kelly M. J. Diederen ![ORCID] https://orcid.org/0000-0002-9619-6583

References

Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. 2013. The computational anatomy of psychosis. Front Psychiatry 4:47.

Aitchison L, Lengyel M. 2017. With or without you: predictive coding and Bayesian inference in the brain. Curr Opin Neurobiol 46:219–27.

Arias-Carrión O, Pŏppel E. 2007. Dopamine, learning, and reward-seeking behavior. Acta Neurobiol Exp (Wars) 67(4):481–8.

Ashby WR. 1952. Design for a brain. Oxford, England: Wiley. ix, 259-ix, 259 p.

Bartra O, McGuire JT, Kable JW. 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76:412–27.

Bassareo V, De Luca MA, Di Chiara G. 2002. Differential expression of motivational stimulus properties by dopamine in nucleus accumbens shell versus core and prefrontal cortex. J Neurosci 22(11):4709–19.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012. Canonical microcircuits for predictive coding. Neuron 76(4):695–711.

Bayer HM, Glimcher PW. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47(1):129–41.

Bayes T, Price R, Canton J. 1763. An essay towards solving a problem in the doctrine of chances. Philos Trans R Soc Lond 53:370–418.

Becker GM, DeGroot MH, Marschak J. 1964. Measuring utility by a single-response sequential method. Behav Sci 9(3):226–32.

Bissonette G, Roesch M. 2016. Development and function of the midbrain dopamine system: what we know and what we need to. Genes Brain Behav 15(1):62–73.

Boorman ED, Rajendran VG, O'Reilly JX, Behrens TE. 2016. Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. Neuron 89(6):1343–54.

Bouret S, Sara SJ. 2004. Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. Eur J Neurosci 20(3):791–802.

Bromberg-Martin ES, Hikosaka O. 2009. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron 63(1):119–26.

Bunzeck N, Düzel E. 2006. Absolute coding of stimulus novelty in the human substantia nigra/VTA. Neuron 51(3):369–79.

Cartoni E, Puglisi-Allegra S, Baldassarre G. 2013. The three principles of action: a Pavlovian-instrumental transfer hypothesis. Front Behav Neurosci 7:153.

Chiodo LA, Antelman SM, Caggiula AR, Lineberry CG. 1980. Sensory stimuli alter the discharge rate of dopamine (DA) neurons: evidence for two functional types of DA cells in the substantia nigra. Brain Res 189(2):544–9.

Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, and others. 2009. Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. Nat Methods 7(2):126–9.

Collins AL, Greenfield VY, Bye JK, Linker KE, Wang AS, Wassum KM. 2016. Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. Sci Rep 6:20231.

Comoli E, Coizet V, Boyes J, Bolam JP, Canteras NS, Quirk RH, and others. 2003. A direct projection from superior colliculus to substantia nigra for detecting salient visual events. Nat Neurosci 6(9):974–80.

Corlett PR, Aitken MR, Dickinson A, Shanks DR, Honey GD, Honey RA, and others. 2004. Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. Neuron 44(5):877–88.

Corlett PR, Honey GD, Aitken MR, Dickinson A, Shanks DR, Absalom AR, and others. 2006. Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. Arch Gen Psychiatry 63(6):611–21

D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. 2008. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 319(5867):1264–7.

Daw ND, Kakade S, Dayan P. 2002a. Opponent interactions between serotonin and dopamine. Neural Netw 15(4):603–16.

Daw ND, Kakade S, Dayan P. 2002b. Opponent interactions between serotonin and dopamine. Neural Netw 15(4–6):603–16.

Deserno L, Boehme R, Heinz A, Schlagenhauf F. 2013. Reinforcement learning and dopamine in schizophrenia: dimensions of symptoms or specific features of a disease group? Front Psychiatry 4:172.

Dickinson A, Balleine B. 2002. The role of learning in the operation of motivational systems. In: Pashler H, Gallistel R, editors. Stevens' handbook of experimental psychology. New York: Wiley. p 497–533.

Düzel E, Bunzeck N, Guitart-Masip M, Düzel S. 2010. NOvelty-related Motivation of Anticipation and exploration by Dopamine (NOMAD): implications for healthy aging. Neurosci Biobehav Rev 34(5):660–9.

de Wit S, Dickinson A. 2009. Associative theories of goal-directed behaviour: a case for animal–human translational models. Psychol Res PRPF 73(4):463–76.

Den Ouden HE, Kok P, De Lange FP. 2012. How prediction errors shape perception, attention, and motivation. Front Psychol 3:548.

Diaconescu AO, Mathys C, Weber LA, Kasper L, Mauer J, Stephan KE. 2017. Hierarchical prediction errors in midbrain and septum during social learning. Soc Cogn Affect Neurosci 12(4):618–34.

Dickinson A, Balleine B. 1994. Motivational control of goal-directed action. Anim Learn Behav 22(1):1–18.

Diederen KMJ, Spencer T, Vestergaard MD, Fletcher PC, Schultz W. 2016. Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. Neuron 90(5):1127–38.

Diederen KMJ, Schultz W. 2015. Scaling prediction errors to reward variability benefits error-driven learning in humans. J Neurophysiol 114(3):1628–40.

Diederen KMJ, Ziauddeen H, Vestergaard MD, Spencer T, Schultz W, Fletcher PC. 2017. Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. J Neurosci 37(7):1708–20.

Dommett E, Coizet V, Blaha CD, Martindale J, Lefebvre V, Walton N, and others. 2005. How visual stimuli activate dopaminergic neurons at short latency. Science 307(5714):1476–9.

Düzel E, Guitart-Masip M, Maass A, Hämmerer D, Betts MJ, Speck O, and others. 2015. Midbrain fMRI: applications, limitations and challenges. fMRI: From Nuclear Spins to Brain Functions: Springer. p 581–609.

Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. 2015. Arithmetic and local circuitry underlying dopamine prediction errors. Nature 525(7568):243–6.

Esber GR, Roesch MR, Bali S, Trageser J, Bissonette GB, Puche AC, and others. 2012. Attention-related Pearce-Kaye-Hall signals in basolateral amygdala require the midbrain dopaminergic system. Biol Psychiatry 72(12):1012–9.

Estes WK. 1948. Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. J Exp Psychol 38(2):173.

Fiorillo CD. 2013. Two dimensions of value: dopamine neurons represent reward but not aversiveness. Science 341(6145):546–9.

Fletcher PC, Anderson JM, Shanks DR, Honey R, Carpenter TA, Donovan T, and others. 2001. Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. Nat Neurosci 4:1043.

Fletcher PC, Frith CD. 2009. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat Rev Neurosci 10(1):48–58.

Fouragnan E, Retzler C, Philiastides MG. 2018. Separate neural representations of prediction error valence and surprise: evidence from an fMRI meta-analysis. Hum Brain Mapp 39(7):2887–906.

Friston K. 2009. The free-energy principle: a rough guide to the brain? Trends Cogn Sci 13(7):293–301.

Friston K. 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci 11(2):127.

Friston KJ, Shiner T, FitzGerald T, Galea JM, Adams R, Brown, and others. 2012. Dopamine, affordance and active inference. PLoS Computational Biol 8(1):e1002327.

Friston KJ, Stephan KE, Montague R, Dolan RJ. 2014. Computational psychiatry: the brain as a phantastic organ. Lancet Psychiatry 1(2):148–58.

Friston KJ, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. 2015. Active inference and epistemic value. Cogn Neurosci 6(4):187–214.

García-García I, Zeighami Y, Dagher A. 2017. Reward prediction errors in drug addiction and Parkinson's disease: from neurophysiology to neuroimaging. Curr Neurol Neurosci Rep 17(6):46.

Garrison J, Erdeniz B, Done J. 2013. Prediction error in RL: a meta-analysis of neuroimaging studies. Neurosci Biobehav Rev 37(7):1297–310.

Gatt JM, Burton KL, Williams LM, Schofield PR. 2015. Specific and common genes implicated across major mental disorders: a review of meta-analysis studies. J Psychiatr Res 60:1–13.

Gläscher J, Daw N, Dayan P, O'Doherty JP. 2010. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free RL. Neuron 66(4):585–95.

Griffiths TL, Chater N, Norris D, Pouget A. 2012. How the Bayesians got their beliefs (and what those beliefs actually are): comment on Bowers and Davis (2012). Psychol Bull 138(3):415–22.

Haarsma J, Fletcher PC, Ziauddeen H, Spencer TJ, Diederen KMJ, Murray GK. 2019. Precision weighting of cortical unsigned prediction errors is mediated by dopamine and benefits learning. bioRxiv.

Halliday GM, Törk I. 1986. Comparative anatomy of the ventromedial mesencephalic tegmentum in the rat, cat, monkey and human. J Comp Neurol 252(4):423–45.

Hart AS, Rutledge RB, Glimcher PW, Phillips PEM. 2014. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J Neurosci 34(3):698–704.

Hayden BY, Heilbronner SR, Pearson JM, Platt ML. 2011. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. J Neurosci 31(11):4178–87.

Heiss W-D. 2009. The potential of PET/MR for brain imaging. Eur J Nucl Med Mol Imaging 36(1):105–12.

Heinz A, Murray GK, Schlagenhauf F, Sterzer P, Grace AA, Waltz JA. 2019. Towards a unifying cognitive, neurophysiological, and computational neuroscience account of schizophrenia. Schizophr Bull 45(5):1092–100.

Howard JD, Kahnt T. 2018. Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. Nat Commun 9(1):1611.

Howes OD, Bose SK, Turkheimer F, Valli I, Egerton A, Valmaggia LR, and others. 2011. Dopamine synthesis capacity before onset of psychosis: a prospective [18F]-DOPA PET imaging study. Am J Psychiatry 168(12):1311–7.

Howes OD, Kambeitz J, Kim E, Stahl D, Slifstein M, Abi-Dargham A, and others. 2012. The nature of dopamine dysfunction in schizophrenia and what this means for treat-ment: meta-analysis of imaging studies. Arch Gen Psychiatry 69(8):776–86.

Howes OD, Murray RM. 2014. Schizophrenia: an integrated sociodevelopmental-cognitive model. Lancet 383(9929):1677–87.

Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HE, and others. 2013. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. Neuron 80(2):519–30.

Jacobs RA, Kruschke JK. 2011. Bayesian learning theory applied to human cognition. Wiley Interdiscip Rev Cogn Sci 2(1):8–21.

Jardri R, Deneve S. 2013. Circular inferences in schizophrenia. Brain 136(11):3227–41.

Jocham G, Klein TA, Ullsperger M. 2014. Differential modulation of RL by D2 dopamine and NMDA glutamate receptor antagonism. J Neurosci 34(39):13151–62.

Kahneman D, Tversky A. 2013. Prospect theory: an analysis of decision under risk. In: MacLean LC, Ziemba WT, editors. Handbook of the fundamentals of financial decision making: Part I. Singapore: World Scientific. p 99–127.

Kakade S, Dayan P. 2002. Dopamine: generalization and bonuses. Neural Netw 15(4–6):549–59.

Kamin LJ. 1969. Predictability, surprise, attention and conditioning. In: Campbell BA, Church RM, editors. Punishment and

aversive behavior. New York: Appleton-Century-Crofts. p 279–96.

Kapur S. 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry 160(1):13–23.

Kempadoo KA, Mosharov EV, Choi SJ, Sulzer D, Kandel ER. 2016. Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. Proc Natl Acad Sci U S A 113(51):14835–40.

Kim KM, Baratta MV, Yang A, Lee D, Boyden ES, Fiorillo CD. 2012. Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. PloS One 7(4):e33612.

Kollins SH, Adcock RA. 2014. ADHD, altered dopamine neurotransmission, and disrupted reinforcement processes: implications for smoking and nicotine dependence. Progress in Neuropsychopharmacol Biol Psychiatry 52:70–8.

Lak A, Stauffer WR, Schultz W. 2014. Dopamine prediction error responses integrate subjective value from different reward dimensions. Proc Natl Acad Sci U S A 111(6):2343–8.

Langdon AJ, Sharpe MJ, Schoenbaum G, Niv Y. 2018. Model-based predictions for dopamine. Curr Opin Neurobiol 49:1–7.

Lapish CC, Shamans JK, Judson Chandler L. 2006. Glutamate-dopamine cotransmission and reward processing in addiction. Alcohol Clin Exp Res 30(9):1451–65.

Levy DJ, Glimcher PW. 2012. The root of all value: a neural common currency for choice. Curr Opin Neurobiol 22(6):1027–38.

Li J, Schiller D, Schoenbaum G, Phelps EA, Daw ND. 2011. Differential roles of human striatum and amygdala in associative learning. Nat Neurosci 14(10):1250–2.

Lisman JE, Grace AA. 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. Neuron 46(5):703–13.

Luce RD. 2012. Individual choice behavior: a theoretical analysis. Chelmsford, MA: Courier Corporation.

Maia TV, Frank MJ. 2017. An integrative perspective on the role of dopamine in schizophrenia. Biol Psychiatry 81(1):52–66.

Mantz J, Thierry A, Glowinski J. 1989. Effect of noxious tail pinch on the discharge rate of mesocortical and mesolimbic dopamine neurons: selective activation of the mesocortical system. Brain Res 476(2):377–81.

Matias S, Lottem E, Dugue GP, Mainen ZF. 2017. Activity patterns of serotonin neurons underlying cognitive flexibility. Elife 6:e20552.

Matsumoto M, Hikosaka O. 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459(7248):837–41.

McClure SM, Berns GS, Montague PR. 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38(2):339–46.

McCutcheon RA, Abi-Dargham A, Howes OD. 2019. Schizophrenia, dopamine and the striatum: from biology to symptoms. Trends Neurosci 42(3):205–20.

Miller R. (1976). Schizophrenic psychology, associative learning and the role of forebrain dopamine. *Medical Hypotheses*, *2*(5):203–211.

Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, and others. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry 13(3):267–76.

Nair-Roberts RG, Chatelain-Badie S, Benson E, White-Cooper H, Bolam J, Ungless M. 2008. Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. Neuroscience 152(4):1024–31.

Nestler EJ, Carlezon Jr WA. 2006. The mesolimbic dopamine reward circuit in depression. Biol Psychiatry 59(12):1151–9.

Niv Y, Schoenbaum G. 2008. Dialogues on prediction errors. Trends Cogn Sci 12(7):265–72.

Nour MM, Dahoun T, Schwartenbeck P, Adams RA, Fitzgerald T, Coello C, and others. 2018. A dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. Proc Natl Acad Sci U S A 115(43):E10167–76.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. 2003. Temporal difference models and reward-related learning in the human brain. Neuron 38(2):329–37.

O'Reilly JX. 2013. Making predictions in a changing world-inference, uncertainty, and learning. Front Neurosci 7:105.

Olds J, Milner P. 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J Comp Physiol Psychol 47(6):419–27.

Papageorgiou GK, Baudonnat M, Cucca F, Walton ME. 2016. Mesolimbic dopamine encodes prediction errors in a state-dependent manner. Cell reports 15(2):221–8.

Padoa-Schioppa C. 2011. Neurobiology of economic choice: a good-based model. Annu Rev Neurosci 34:333–59.

Pavlov IP, Anrep GV. 1928. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. London: Oxford University Press.

Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP. 2013. The neural representation of unexpected uncertainty during value-based decision making. Neuron 79(1):191–201.

Pearce JM, Hall G. 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol Rev 87(6):532–52.

Pennartz CMA, McNaughton BL, Mulder AB. 2000. The glutamate hypothesis of reinforcement learning. Prog Brain Res 126:231–53.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442(7106):1042–5.

Phillips PEM, Stuber GD, Heien ML, Wightman RM, Carelli RM. 2003. Subsecond dopamine release promotes cocaine seeking. Nature 422(6932):614–8.

Radua J, Schmidt A, Borgwardt S, Heinz A, Schlagenhauf, McGuire, and others. 2015. Ventral striatal activation during reward processing in psychosis: a neurofunctional meta-analysis. JAMA Psychiatry 72(12):1243–51.

Rangel-Gomez M, Meeter M. 2016. Neurotransmitters and novelty: a systematic review. J Psychopharmacol 30(1):3–12.

Reed P, Mitchell C, Nokes T. 1996. Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. Anim Learn Behav 24(1):38–45.

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. Classical conditioning II: current research and theory. New York: Appleton Century Crofts. p 64–99.

Robinson TE, Berridge KC. 2008. The incentive sensitization theory of addiction: some current issues. Philos Trans R Soc Lond B Biol Sci 363(1507):3137–46.

Roiser J, Stephan K, Den Ouden H, Barnes T, Friston K, Joyce E. 2009. Do patients with schizophrenia exhibit aberrant salience? Psychol Med 39(2):199–209.

Schultz W. 1998. Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.

Schultz W. 2016a. Dopamine reward prediction error coding. Dialogues Clin Neurosci 18(1):23–32.

Schultz W. 2016b. Dopamine reward prediction-error signalling: a two-component response. Nat Rev Neurosci 17(3):183.

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. Science 275(5306):1593–9.

Schultz W, Dickinson A. 2000. Neuronal coding of prediction errors. Annu Rev Neurosci 23:473–500.

Schwartenbeck P, FitzGerald THB, Dolan R. 2016. Neural signals encoding shifts in beliefs. Neuroimage 125:578–86.

Sescousse G, Caldú X, Segura B, Dreher J-C. 2013. Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci Biobehav Rev 37(4):681–96.

Sharpe MJ, Chang CY, Liu MA, Batchelor HM, Mueller LE, Jones JL, and others. 2017. Dopamine transients are sufficient and necessary for acquisition of model-based associations. Nat Neurosci 20:735.

Sharpe MJ, Marchant NJ, Whitaker LR, Richie CT, Zhang YJ, Campbell EJ, and others 2017. Lateral hypothalamic GABAergic neurons encode reward predictions that are relayed to the ventral tegmental area to regulate learning. Curr Biol 27(14):2089–100.

Skinner B. 1938. The behavior of organisms: an experimental analysis. New York: Appleton-Century.

Skinner BF. 1963. Operant behavior. Am Psychol 18(8):503.

Solomon RL, Corbit JD. 1978. An opponent-process theory of motivation. Am Econ Rev 68(6):12–24.

Starkweather CK, Babayan BM, Uchida N, Gershman SJ. 2017. Dopamine reward prediction errors reflect hidden-state inference across time. Nat Neurosci 20(4):581.

Stauffer WR, Lak A, Bossaerts P, Schultz W. 2015. Economic choices reveal probability distortion in macaque monkeys. J Neurosci 35(7):3146–54.

Stauffer WR, Lak A, Yang A, Borel M, Paulsen O, Boyden ES, and others. 2016. Dopamine neuron-specific optogenetic stimulation in rhesus macaques. Cell 166(6):1564–71.e6.

Sterzer P, Adams RA, Fletcher P, Frith C, Lawrie SM, Muckli L, and others. 2018. The predictive coding account of psychosis. Biol Psychiatry 84(9):634–43.

Suri RE, Bargas J, Arbib MA. 2001. Modeling functions of striatal dopamine modulation in learning and planning. Neuroscience 103(1):65–85.

Sutton RS, Barto AG. 1998. Reinforcement learning: an introduction. Cambridge, MA: MIT Press.

Takahashi YK, Batchelor HM, Liu B, Khanna A, Morales M, Schoenbaum G. 2017. Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. Neuron 95(6):1395–405.

Taylor M, Creelman CD. 1967. PEST: efficient estimates on probability functions. J Acoust Soc Am 41(4A):782–7.

Thorndike EL. 1898. Animal intelligence: An experimental study of the associative processes in animals. Psychol Rev Monogr Suppl 2(4):i–109.

Tobler PN, Christopoulos GI, O'Doherty JP, Dolan RJ, Schultz W. 2008. Neuronal distortions of reward probability without choice. J Neurosci 28(45):11703–11.

Tobler PN, Dickinson A, Schultz W. 2003. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. J Neurosci 23(32):10402–10.

Tobler PN, Fiorillo CD, Schultz W. 2005. Adaptive coding of reward value by dopamine neurons. Science 307(5715): 1642–5.

Turner DC, Aitken MRF, Shanks DR, Sahakian BJ, Robbins TW, Schwarzbauer C, and others. 2004. The role of the lateral frontal cortex in causal associative learning: exploring preventative and super-learning. Cereb Cortex 14(8):872–80.

Tversky A, Kahneman D. 1974. Judgment under uncertainty: Heuristics and biases. Science 185(4157):1124–31.

Ungless MA. 2004. Dopamine: the salient issue. Trends Neurosci 27(12):702–6.

Ungless MA, Magill PJ, Bolam JP. 2004. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science 303(5666):2040–2.

Von Helmholtz H. 1867. Handbuch der physiologischen Optik. Leipzig, Germany: Leopold Voss.

Valton V, Romaniuk L, Steele JD, Lawrie S, Seriès P. 2017. Comprehensive review: computational modelling of schizophrenia. Neurosci Biobehav Rev 83:631–46.

Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, and others. 2018. Prefrontal cortex as a meta-RL system. Nat Neurosci 21(6):860–8.

Watabe-Uchida M, Eshel N, Uchida N. 2017. Neural circuitry of reward prediction error. Annu Rev Neurosci 40(1):373–94.

White DM, Kraguljac NV, Reid MA, Lahti AC. 2015. Contribution of substantia nigra glutamate to prediction error signals in schizophrenia: a combined magnetic resonance spectroscopy/functional imaging study. NPJ Schizophr 1(1):1–7.

Wickens JR, Reynolds JN, Hyland BI. 2003. Neural mechanisms of reward-related motor learning. Curr Opin Neurobiol 13(6):685–90.

Wise RA. 2002. Brain reward circuitry: insights from unsensed incentives. Neuron 36(2):229–40.

Wunderlich K, Smittenaar P, Dolan RJ. 2012. Dopamine enhances model-based over model-free choice behavior. Neuron 75(3):418–24.

Yamada H, Tymula A, Louie K, Glimcher PW. 2013. Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. Proc Natl Acad Sci U S A 110(39): 15788–93.

Yu AJ, Dayan P. 2005. Uncertainty, neuromodulation, and attention. Neuron 46(4):681–92.