HHS Public Access

What does dopamine mean?

Josh Berke

Departments of Neurology and Psychiatry, and Kavli Institute for Fundamental Neuroscience, University of California, San Francisco, San Francisco, CA, USA. joshua.berke@ucsf.edu.

Abstract

Dopamine is a critical modulator of both learning and motivation. This presents a problem: how can target cells know whether increased dopamine is a signal to learn, or to move? It is often presumed that motivation involves slow ("tonic") dopamine changes, while fast ("phasic") dopamine fluctuations convey reward prediction errors for learning. Yet recent studies have shown that dopamine conveys motivational value, and promotes movement, even on sub-second timescales. Here I describe an alternative account of how dopamine regulates ongoing behavior. Dopamine release related to motivation is rapidly and locally sculpted by receptors on dopamine terminals, independently from dopamine cell firing. Target neurons abruptly switch between learning and performance modes, with striatal cholinergic interneurons providing one candidate switch mechanism. The behavioral impact of dopamine varies by subregion, but in each case dopamine provides a dynamic estimate of whether it is worth expending a limited internal resource, such as energy, attention, or time.

Is dopamine a signal for learning, for motivation, or both?

Our understanding of dopamine has changed in the past, and is changing once again. One critical distinction is between dopamine effects on <u>current</u> behavior (performance), and dopamine effects on <u>future</u> behavior (learning). Both are real and important, but at various times one has been in favor and the other has not.

When (in the '70s) it became possible to perform selective, complete lesions of dopamine pathways, the obvious behavioral consequence was a severe reduction in movement[1]. This fit with the akinetic effects of dopamine loss in humans, produced by advanced Parkinson's disease, toxic drugs, or encephalitis[2]. Yet neither rat nor human cases display a fundamental inability to move. Dopamine-lesioned rats swim in cold water[3], and akinetic patients may get up and run if a fire alarm sounds("paradoxical" kinesia). Nor is there a basic deficit in appreciating rewards: dopamine-lesioned rats will consume food placed in their mouths, and show signs of enjoying it[4]. Rather, they will not choose to exert effort to actively obtain rewards. These and many other results established a fundamental link between dopamine and motivation[5]. Even the movement slowing observed in less-severe cases of Parkinson's Disease can be considered a motivational deficit, reflecting implicit decisions that it is not worth expending the energy required for faster movements[6].

Then (in the '80s) came pioneering recordings of dopamine neurons in behaving monkeys(in midbrain areas that project to forebrain: ventral tegmental area, VTA / substantia nigra pars compacta, SNc). Among observed firing patterns were brief bursts of activity to stimuli that

triggered immediate movements. This "phasic" dopamine firing was initially interpreted as supporting "behavioral activation"[7] and "motivational arousal"[8] - in other words, as invigorating the animal's current behavior.

A radical shift occurred in the '90s, with the reinterpretation of phasic dopamine bursts as encoding *reward prediction errors* (RPEs[9]). This was based upon a key observation: dopamine cells respond to unexpected stimuli associated with future reward, but often stop responding if these stimuli become expected[10]. The RPE idea originated in earlier learning theories, and especially in the then-developing computer science field of reinforcement learning[11]. The point of an RPE signal is to update *values* (estimates of future rewards). These values are used later, to help make choices that maximize reward. Since dopamine cell firing resembled RPEs, and RPEs are used for learning, it became natural to emphasize the role of dopamine in learning. Later optogenetic manipulations confirmed the dopaminergic identity of RPE-coding cells[12,13] and showed they indeed modulate learning[14,15].

The idea that dopamine provides a learning signal fits beautifully with the literature that dopamine modulates synaptic plasticity in the striatum, the primary forebrain target of dopamine. For example, the triple coincidence of glutamate stimulation of a striatal dendrite spine, postsynaptic depolarization, and dopamine release causes the spine to grow[16]. Dopaminergic modulation of long-term learning mechanisms helps explain the persistent behavioral effects of addictive drugs, which share the property of enhancing striatal dopamine release[17]. Even the profound akinesia with dopamine loss can be partly accounted for by such learning mechanisms[18]. Lack of dopamine may be treated as a constantly-negative RPE, that progressively updates values of actions towards zero. Similar progressive, extinction-like effects on behavior can be produced by dopamine antagonists[19,20].

Yet the idea that dopamine is critically involved in ongoing motivation has never gone away - on the contrary, it is widely taken for granted by behavioral neuroscientists. This is appropriate given the strong evidence that dopamine functions in motivation/movement/invigoration are dissociable from learning[15,20–23]. Less widely-appreciated is the challenge involved in reconciling this motivational role with the theory that DA provides an RPE learning signal.

Motivation "looks forward": it uses predictions of future reward (values) to appropriately energize current behavior. By contrast, learning "looks backwards" at states and actions in the recent past, and updates their values. These are complementary phases of a cycle: the updated values may be used in subsequent decision-making if those states are re-encountered, then updated again, and so forth. But which phase of the cycle is dopamine involved with – using values to make decisions (performance), or updating values (learning)?

In some circumstances it is straightforward to imagine dopamine playing both roles simultaneously.[24] Unexpected, reward-predictive cues are the archetypal events for evoking dopamine cell firing and release, and such cues typically both invigorate behavior and evoke learning (Fig. 1). In this particular situation both reward prediction, and reward prediction errors, increase simultaneously - but this is not always the case. As just one example, people

and other animals are often motivated to work for rewards even when little or nothing surprising occurs. They may work harder and harder as they get closer and closer to reward (value increases as rewards draw near). The point is that learning and motivation are conceptually, computationally, and behaviorally distinct - and yet dopamine seems to do both.

Below I critically assess current ideas about how dopamine is able to achieve both learning and motivational functions. I propose an updated model, based on three key facts: 1) dopamine release from terminals does not arise simply from dopamine cell firing, but can also be locally controlled; 2) dopamine affects both synaptic plasticity and excitability of target cells, with distinct consequences for learning and performance respectively; 3) dopamine effects on plasticity can be switched on or off by nearby circuit elements. Together, these features may allow brain circuits to toggle between two distinct dopamine messages, for learning and motivation respectively.

Are there separate "phasic" and "tonic" dopamine signals, with different meanings?

It is often argued that the learning and motivational roles of dopamine occur on different time scales[25]. Dopamine cells fire continuously ("tonically") at a few spikes per second, with occasional brief ("phasic") bursts or pauses. Bursts, especially if artificially synchronized across dopamine cells, drive corresponding rapid increases in forebrain dopamine[26] that are highly transient (sub-second duration[27]). The separate contribution of tonic dopamine cell firing to forebrain dopamine concentrations is less clear. Some evidence suggests this contribution is very small[28]. It may be sufficient to produce near-continuous stimulation of the higher-affinity D2 receptors, allowing the system to notice brief pauses in dopamine cell firing[29] and use these pauses as negative prediction errors.

Microdialysis has been widely used to directly measure forebrain dopamine levels, albeit with low temporal resolution(typically averaging across many minutes). Such slow measurements of dopamine can be challenging to relate precisely to behavior. Nonetheless microdialysis of dopamine in the nucleus accumbens(NAc; ventral/medial striatum) shows positive correlations to locomotor activity[30] and other indices of motivation[5]. This has been widely taken to mean that there are slow("tonic") changes in dopamine concentration, and that these slow changes convey a motivational signal. More specifically, computational models have proposed that tonic dopamine levels track the long-term average reward rate[31] - a useful motivational variable for time allocation and foraging decisions. It is worth emphasizing that very few papers clearly define "tonic" dopamine levels – they usually just assume that dopamine concentration slowly changes over the multiple-minutes time scale of microdialysis.

Yet this "phasic dopamine=RPE/learning, tonic dopamine=motivation" view faces many problems. First, there is no direct evidence that tonic dopamine cell firing normally varies over slow time scales. Tonic firing rates do not change with changing motivation[32,33]. It has been argued that tonic dopamine levels change due to a changing proportion of active

dopamine cells[34,35]. But across many studies in undrugged, unlesioned animals, dopamine cells have never been reported to switch between silent and active states.

Furthermore, the fact that microdialysis measures dopamine levels slowly does not mean that dopamine levels actually change slowly. We recently[15] examined rat NAc dopamine in a probabilistic reward task, using both microdialysis and fast-scan cyclic voltammetry. We confirmed that mesolimbic dopamine, as measured by microdialysis, correlates with reward rate(rewards/min). However, even with an improved microdialysis temporal resolution(1min) dopamine fluctuated as fast as we sampled it: we saw no evidence for an inherently-slow dopamine signal.

Using the finer temporal resolution still of voltammetry we observed a close relationship between sub-second dopamine fluctuations and motivation. As rats performed the sequence of actions needed to achieve rewards, dopamine rose higher and higher, reaching a peak just as they obtained the reward(and dropping rapidly as they consumed it). We showed that dopamine correlated strongly with instantaneous state value -defined as the expected future reward, discounted by the expected time needed to receive it. These rapid dopamine dynamics can also explain the microdialysis results, without invoking separate dopamine signals on different time scales. As animals experience more rewards, they increase their expectations of future rewards at each step in the trial sequence. Rather than a slowly-evolving average reward rate signal, the correlation between dopamine and reward rate is best explained as an average, over the prolonged microdialysis sample collection time, of these rapidly-evolving state values.

This value interpretation of mesolimbic dopamine release is consistent with voltammetry results from other research groups, who have repeatedly found that dopamine release ramps up with increasing proximity to reward[36–38](Fig. 2). This motivational signal is not inherently "slow", but rather can be observed across a continuous range of time scales. Although dopamine ramps can last several seconds when an approach behavior also lasts several seconds[38], this reflects the time course of the behavior, rather than intrinsic dopamine dynamics. The relationship between mesolimbic dopamine release and fluctuating value is visible as fast as the recording technique permits, i.e. on a ~100ms timescale with acute voltammetry electrodes[15].

Fast dopamine fluctuations do not simply mirror motivation, they also immediately drive motivated behavior. Larger phasic responses of dopamine cells to trigger cues predict shorter reaction times on that very same trial[39]. Optogenetic stimulation of VTA dopamine cells makes rats more likely to begin work in our probabilistic reward task[15], just as if they had a higher expectation of reward. Optogenetic stimulation of SNc dopamine neurons, or their axons in dorsal striatum, increases the probability of movement[40,41]. Critically, these behavioral effects are apparent within a couple hundred milliseconds of the onset of optogenetic stimulation. The ability of reward-predictive cues to boost motivation appears to be mediated by very rapid dopaminergic modulation of the excitability of NAc spiny neurons[42]. Since dopamine is changing quickly, and these dopamine changes affect motivation quickly, the motivational functions of dopamine are better described as fast("phasic"), not slow("tonic").

Furthermore, invoking separate fast and slow time scales does not in itself solve the decoding problem faced by neurons with dopamine receptors. If dopamine signals learning, modulation of synaptic plasticity would seem an appropriate cellular response. But immediate effects on motivated behavior imply immediate effects on spiking - e.g. through rapid changes in excitability. Dopamine can have both of these postsynaptic effects(and more), so does a given dopamine concentration have a specific meaning? Or does this meaning need to be constructed - e.g. by comparing dopamine levels across time, or by using other coincident signals to determine which cellular machinery to engage? This possibility is discussed further below.

Does dopamine release convey the same information as dopamine cell firing?

The relationship between fast dopamine fluctuations and motivational value seems strange, given that dopamine cell firing instead resembles RPE. Furthermore, some studies have reported RPE signals in mesolimbic dopamine release[43]. It is important to note a challenge in interpreting some forms of neural data. Value signals and RPEs are correlated with each other - not surprisingly as the RPE is usually defined as the change in value from one moment to the next("temporal-difference" RPE). Because of this correlation it is critical to use experimental designs and analyses that distinguish value from RPE accounts. The problem is compounded when using a neural measure that relies on relative, rather than absolute, signal changes. Voltammetry analyses usually compare dopamine at some time point of interest to a "baseline" epoch earlier in each trial(to remove signal components that are non-dopamine-dependent, including electrode charging on each voltage sweep and drift over a timescale of minutes). But subtracting away a baseline can make a value signal resemble an RPE signal. This is what we observed in our own voltammetry data(Fig. 2e). Changes in reward expectation were reflected in changes in dopamine concentration early in each trial, and these changes are missed if one just assumes a constant baseline across trials[15]. Conclusions about dopamine release and RPE coding thus need to be viewed with caution. This data interpretation danger applies not only to voltammetry, but to any analysis that relies on relative changes - potentially including some fMRI and photometry[44].

Nonetheless, we still need to reconcile value-related dopamine release in NAc core with the consistent absence of value-related spiking by dopamine neurons[13], even within the lateral VTA area that provides dopamine to NAc core[45]. One potential factor is that dopamine cells are usually recorded in head-restrained animals performing classical conditioning tasks, while dopamine release is typically measured in unrestrained animals actively moving through their environment. We proposed that mesolimbic dopamine might specifically indicate the value of "work"[15] - that it reflects a requirement for devoting time and effort to obtain the reward. Consistent with this, dopamine increases with signals instructing movement, but not with signals instructing stillness, even when they indicate similar future reward[46]. If - as in many classical conditioning tasks - there is no benefit to active "work", then dopaminergic changes indicating the value of work may be less apparent.

Even more important may be the fact that dopamine release can be locally controlled at the terminals themselves, and thus show spatio-temporal patterns independent of cell body spiking. For example, the basolateral amygdala (BLA) can influence NAc dopamine release even when VTA is inactivated[47]. Conversely, inactivating BLA reduces NAc dopamine release and corresponding motivated behavior, without apparently affecting VTA firing[48]. Dopamine terminals have receptors for a range of neurotransmitters, including glutamate, opioids, and acetylcholine. Nicotinic acetylcholine receptors allow striatal cholinergic interneurons(CINs) to rapidly control dopamine release[49,50]. Although it has long been noted that local control of dopamine release is potentially important[7,51], it has not been included in computational accounts of dopamine function. I propose that dopamine release dynamics related to value coding arise largely through *local* control, even as dopamine cell firing provides important RPE-like signals for learning.

How can dopamine mean both learning and motivation without confusion?

In principle, a value signal is sufficient to convey RPE as well, since temporal-difference RPEs simply are rapid changes in value(Fig. 2B). For example, distinct intracellular pathways in target neurons might be differently sensitive to the absolute concentration of dopamine(representing value) versus fast relative changes in concentration(representing RPE). This scheme seems plausible, given the complex dopamine modulation of spiny neuron physiology[52] and their sensitivity to temporal patterns of calcium concentration[53]. Yet this also seems somewhat redundant. If an RPE-like signal already exists in dopamine cell spiking, it ought to be possible to use it rather than re-deriving RPE from a value signal.

To appropriately use distinct RPE and value signals, dopamine-recipient circuits may actively switch how they interpret dopamine. There is intriguing evidence that acetylcholine may serve this switching role too. At the same time as dopamine cells fire bursts of spikes to unexpected cues, CINs show brief(~150ms) *pauses* in firing, which do not scale with RPEs[54]. These CIN pauses can be driven by VTA GABAergic neurons[55] as well as "surprise"-related cells in the intralaminar thalamus, and have been proposed to act as an associability signal promoting learning[56]. Morris and Bergman suggested[54] that cholinergic pauses define temporal windows for striatal plasticity, during which dopamine can be used as a learning signal. Dopamine-dependent plasticity is continuously suppressed by mechanisms including muscarinic m4 receptors on direct-pathway striatal neurons[57]. Models of intracellular signaling suggest that during CIN pauses, the absence of m4 binding may act synergistically with phasic dopamine bursts to boost PKA activation[58], thereby promoting synaptic change.

Striatal cholinergic cells are thus well-positioned to dynamically switch the meaning of a multiplexed dopaminergic message. During CIN pauses, relief of a muscarinic block over synaptic plasticity would allow dopamine to be used for learning. At other times release from dopamine terminals would be locally sculpted to affect ongoing behavioral performance. Currently, this suggestion is both speculative and incomplete. It has been proposed that CINs integrate information from many surrounding spiny neurons to extract useful network-level signals such as entropy[59,60]. But it is not at all clear that CIN activity

dynamics can be used to generate dopamine value signals[61], and also to gate dopamine learning signals.

Does dopamine mean the same thing throughout the forebrain?

As the RPE idea took hold, it was imagined that dopamine was a global signal, broadcasting an error message throughout striatal and frontal cortical targets. Schultz emphasized that monkey dopamine cells throughout VTA and SNc have very similar responses[62]. Studies of identified dopamine cells have also found quite homogeneous RPE-like responses in rodents, at least for lateral VTA neurons within classical conditioning contexts[13]. Yet dopamine cells are molecularly and physiologically diverse[63–65] and there are now many reports that they show diverse firing patterns in behaving animals. These include phasic increases in firing to aversive events[66] and trigger cues[67] that fit poorly with the standard RPE account. Many dopamine cells show an initial short-latency response to sensory events that reflects surprise or "alerting" more than specific RPE coding[68,69]. This alerting aspect is more prominent in SNc[69], where dopamine cells project more to "sensorimotor" dorsal/lateral striatum(DLS[45,63]). Subpopulations of SNc dopamine cells have also been reported to increase[41] or decrease[70] firing in conjunction with spontaneous movements, even without external cues.

Several groups used fiber photometry and the calcium indicator GCaMP to examine bulk activity of subpopulations of dopamine neurons[71,72]. Dopamine cells that project to the dorsal/medial striatum(DMS) showed transiently depressed activity to unexpected brief shocks, while those projecting to DLS showed increased activity[71]- more consistent with an alerting response. Distinct dopaminergic responses in different forebrain subregions have also been observed using GCaMP to examine activity of dopamine axons and terminals[40,72,73]. Using two-photon imaging in head-restrained mice, Howe and Dombeck[40] reported phasic dopamine activity related to spontaneous movements. This was predominantly seen in individual dopamine axons from SNc that terminated in dorsal striatum, while VTA dopamine axons in NAc responded more to reward delivery. Others also found reward-related dopaminergic activity in NAc, with DMS instead more linked to contralateral actions[72] and the posterior tail of striatum responsive to aversive and novel stimuli[74].

Direct measures of dopamine release also reveal heterogeneity between subregions[30,75]. With microdialysis we found dopamine to be correlated with value specifically in NAc core and ventral-medial frontal cortex, not in other medial parts of striatum(NAc shell, DMS) or frontal cortex. This is intriguing as it appears to map well to two "hotspots" of value coding consistently seen in human fMRI studies studies[76,77]. In particular the NAc BOLD signal, which has a close relationship to dopamine signaling[78], increases with reward anticipation(value) - more than with RPE[76].

Whether these spatial patterns of dopamine release arise from firing of distinct dopamine cell subpopulations, local control of dopamine release, or both, they challenge the idea of a global dopamine message. One might conclude that there are many different dopamine functions, with (for example) dopamine in dorsal striatum signaling "movement" and

dopamine in ventral striatum signaling "reward"[40]. However, I favor another conceptual approach. Different striatal subregions get inputs from different cortical regions, and so will be processing different types of information. Yet each striatal subregion shares a common microcircuit architecture, including separate D1- versus D2- receptor bearing spiny neurons[79], CINs, and so forth. Although it is common to refer to various striatal subregions(e.g. DLS, DMS, NAc core) as if they are discrete areas, there are no sharp anatomical boundaries between them(NAc shell is a bit more neurochemically distinct). Instead there are just gentle gradients in receptor density, interneuron proportions etc., which seem more like tweaks to the parameters of a shared computational algorithm. Given this common architecture, can we describe a common dopamine function, abstracted away from the specific information being handled by each subregion?

Striatal dopamine and the allocation of limited resources.

I propose that a variety of disparate dopamine effects on ongoing behavior can be understood as modulation of *resource allocation decisions*. Specifically, dopamine provides estimates of how worthwhile it is to expend a limited internal resource, with the particular resource differing between striatal subregions. For "motor" striatum(~DLS) the resource is movement, which is limited because moving costs energy, and because many actions are incompatible with each other[80]. Increasing dopamine makes it more likely that an animal will decide it is worth expending energy to move, or move faster[6,40,81]. Note that a dopamine signal that encodes "movement is worthwhile" will produce correlations between dopamine and movement, even without dopamine encoding "movement" *per se*.

For "cognitive" striatum(~DMS) the resources are cognitive processes including attention(which is limited-capacity by definition[82]) and working memory[83]. Without dopamine, salient external cues that normally provoke orienting movements are neglected, as if considered less attention-worthy[3]. Furthermore, deliberately marshaling cognitive control processes is effortful(costly[84]). Dopamine – especially in DMS[85] - plays a key role in deciding whether it is worth exerting this effort[86,87]. This can include whether to employ more cognitively-demanding, deliberative("model-based") decision strategies[88].

For "motivational" striatum(~NAc) one key limited resource may be the animal's time. Mesolimbic dopamine is not required when animals perform a simple, fixed action to rapidly obtain rewards[89]. But many forms of reward can only be obtained through prolonged work: extended sequences of unrewarded actions, as in foraging. Choosing to engage in work means that other beneficial ways of spending time must be foregone. High mesolimbic dopamine indicates that engaging in temporally-extended, effortful work is worthwhile, but as dopamine is lowered animals do not bother, and may instead just prepare to sleep[90].

Within each cortico-striatal loop circuit dopamine's contribution to ongoing behavior is thus both economic(concerned with resource allocation) and motivational(whether it is *worthwhile* to expend resources[81]). These circuits are not fully independent, but rather have a hierarchical, spiraling organization: more ventral portions of striatum influence dopamine cells that project to more dorsal portions[5,91]. In this way decisions to engage in work may also help invigorate required specific, briefer movements. But overall, dopamine provides

"activational" signals - increasing the probability that some decision is made - rather than "directional" signals specifying *how* resources should be spent[5].

What is the computational role of dopamine as decisions are made?

One way of thinking about this activational role is in terms of decision-making "thresholds". In certain mathematical models decision processes increase until they reach a threshold level, when the system becomes committed to an action[92]. Higher dopamine would be equivalent to a lower distance-to-threshold, so that decisions are reached more rapidly. This idea is simplistic, yet makes quantitative predictions that have been confirmed. Lowering thresholds for movement would cause a specific change in the shape of the reaction time distribution, just what is seen when amphetamine is infused into sensorimotor striatum[20].

Rather than fixed thresholds, behavioral and neural data may be better fit if thresholds decrease over time, as if decisions become increasingly urgent. Basal ganglia output has been proposed to provide a dynamically-evolving urgency signal, which invigorates selection mechanisms in cortex[93]. Urgency was also greater when future rewards were closer in time, making this concept similar to the value coding, activational role of dopamine.

Is such an activational role sufficient to describe the performance-modulating effects of striatal dopamine? This is related to the long-standing question of whether basal ganglia circuits directly select among learned actions[80] or merely invigorate choices made elsewhere[93,94]. There are at least two ways in which dopamine can appear to have a more "directional" effect. The first is when dopamine acts within a brain subregion that processes inherently directional information. Basal ganglia circuits have an important, partly-lateralized role orienting towards and approaching potential rewards. The primate caudate(~DMS) is involved in driving eye movements towards contralateral spatial fields[95]. A dopaminergic signal that something in contralateral space is worth orienting towards may account for the observed correlation between dopaminergic activity in DMS and contralateral movements[72], as well as the rotational behavior produced by dopamine manipulations[96]. A second "directional" influence of dopamine is apparent when (bilateral) dopamine lesions bias rats towards low-effort / low-reward choices, rather than high-effort / high-reward alternatives[97]. This may reflect the fact that some decisions are more serial than parallel, with rats(and humans) evaluating options one-at-a-time[98]. In these decision contexts dopamine may still pay a fundamentally activational role by conveying the value of the currently-considered option, which can then be accepted or not[24].

Active animals make decisions at multiple levels, often at high rates. Beyond thinking about individual decisions, it may be helpful to consider an overall trajectory through a sequence of states(Fig. 1). By facilitating transitions from one state to the next, dopamine may accelerate the flow along learned trajectories[99]. This may relate to the important influence of dopamine over the timing of behavior[44,100]. One key frontier for future work is to gain a deeper understanding of how such dopamine effects on ongoing behavior arise mechanistically, by altering information processing within single cells, microcircuits and large-scale cortical-basal ganglia loops. Also, I have emphasized common computational roles of dopamine across a range of striatal targets, but largely neglected cortical targets, and

it remains to be seen whether dopamine functions in both structures can be described within the same framework.

In summary, an adequate description of dopamine would explain how dopamine can signal both learning, and motivation, on the same fast time scales, without confusion. It would explain why dopamine release in key targets covaries with reward expectation even though dopamine cell firing does not. And it would provide a unified computational account of dopamine actions throughout striatum and elsewhere, which explains disparate behavioral effects on movement, cognition, and timing. Some specific ideas presented here are speculative, but are intended to invigorate renewed discussion, modeling, and incisive new experiments.

Acknowledgements.

References:

1. Ungerstedt U Adipsia and aphagia after 6-hydroxydopamine induced degeneration of the nigro-striatal dopamine system. Acta Physiol Scand Suppl 367, 95–122 (1971). [PubMed: 4332694]

2. Sacks O Awakenings. Awakenings (1973).

3. Marshall JF, Levitan D and Stricker EM Activation-induced restoration of sensorimotor functions in rats with dopamine-depleting brain lesions. J Comp Physiol Psychol 90, 536–46 (1976). [PubMed: 8470]

4. Berridge KC, Venier IL and Robinson TE Taste reactivity analysis of 6-hydroxydopamine-induced aphagia: implications for arousal and anhedonia hypotheses of dopamine function. Behav Neurosci 103, 36–45 (1989). [PubMed: 2493791]

5. Salamone J and Correa M The Mysterious Motivational Functions of Mesolimbic Dopamine. Neuron 76, 470–485 (2012).doi:10.1016/j.neuron.2012.10.021 [PubMed: 23141060]

6. Mazzoni P, Hristova A and Krakauer JW Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. J Neurosci 27, 7105–16 (2007).doi:10.1523/JNEUROSCI.0264-07.2007 [PubMed: 17611263]

7. Schultz W Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. Journal of neurophysiology 56, 1439–1461 (1986). [PubMed: 3794777]

8. Schultz W and Romo R Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. J Neurophysiol 63, 607–24 (1990). [PubMed: 2329364]

9. Montague PR, Dayan P and Sejnowski TJ A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16, 1936–47 (1996). [PubMed: 8774460]

10. Schultz W, Apicella P and Ljungberg T Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J Neurosci 13, 900–13 (1993). [PubMed: 8441015]

11. Sutton RS and Barto AG Reinforcement learning: an introduction. Reinforcement learning: an introduction (MIT Press: Cambridge, Massachusetts, 1998).

12. Cohen JY, Haesler S, Vong L, Lowell BB and Uchida N Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482, 85–8 (2012).doi:10.1038/nature10754 [PubMed: 22258508]

13. Eshel N, Tian J, Bukwich M and Uchida N Dopamine neurons share common response function for reward prediction error. Nat Neurosci 19, 479–86 (2016).doi:10.1038/nn.4239 [PubMed: 26854803]

14. Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K and Janak PH A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci (2013).doi:10.1038/nn.3413

15. Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ and Berke JD Mesolimbic dopamine signals the value of work. Nat Neurosci 19, 117–26 (2016).doi:10.1038/nn.4173 [PubMed: 26595651]

16. Yagishita S, Hayashi-Takagi A, Ellis-Davies GC, Urakubo H, Ishii S and Kasai H A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science 345, 1616–20 (2014).doi:10.1126/science.1255514 [PubMed: 25258080]

17. Berke JD and Hyman SE Addiction, dopamine, and the molecular mechanisms of memory. Neuron 25, 515–32 (2000). [PubMed: 10774721]

18. Beeler JA, Frank MJ, McDaid J, Alexander E, Turkson S, Bernandez MS, McGehee DS and Zhuang X A role for dopamine-mediated learning in the pathophysiology and treatment of Parkinson's disease. Cell Rep 2, 1747–61 (2012).doi:10.1016/j.celrep.2012.11.014 [PubMed: 23246005]

19. Wise RA Dopamine, learning and motivation. Nat Rev Neurosci 5, 483–94 (2004).doi:10.1038/nrn1406 [PubMed: 15152198]

20. Leventhal DK, Stoetzner C, Abraham R, Pettibone J, DeMarco K and Berke JD Dissociable effects of dopamine on learning and performance within sensorimotor striatum. Basal Ganglia 4, 43–54 (2014).doi:10.1016/j.baga.2013.11.001 [PubMed: 24949283]

21. Wyvell CL and Berridge KC Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. J Neurosci 20, 8122–30. (2000). [PubMed: 11050134]

22. Cagniard B, Beeler JA, Britt JP, McGehee DS, Marinelli M and Zhuang X Dopamine scales performance in the absence of new learning. Neuron 51, 541–7 (2006).doi:10.1016/j.neuron.2006.07.026 [PubMed: 16950153]

23. Shiner T, Seymour B, Wunderlich K, Hill C, Bhatia KP, Dayan P and Dolan RJ Dopamine and performance in a reinforcement learning task: evidence from Parkinsons disease. Brain 135, 1871–1883 (2012). [PubMed: 22508958]

24. McClure SM, Daw ND and Montague PR A computational substrate for incentive salience. Trends Neurosci 26, 423–8 (2003). [PubMed: 12900173]

25. Schultz W Multiple dopamine functions at different time courses. Annu Rev Neurosci 30, 259–88 (2007).doi:10.1146/annurev.neuro.28.061604.135722 [PubMed: 17600522]

26. Gonon F, Burie JB, Jaber M, Benoit-Marand M, Dumartin B and Bloch B Geometry and kinetics of dopaminergic transmission in the rat striatum and in mice lacking the dopamine transporter. Prog Brain Res 125, 291–302 (2000). [PubMed: 11098665]

27. Aragona BJ, Cleaveland NA, Stuber GD, Day JJ, Carelli RM and Wightman RM Preferential enhancement of dopamine transmission within the nucleus accumbens shell by cocaine is attributable to a direct increase in phasic dopamine release events. J Neurosci 28, 8821–31 (2008).doi:10.1523/JNEUROSCI.2225-08.2008 [PubMed: 18753384]

28. Owesson-White CA, Roitman MF, Sombers LA, Belle AM, Keithley RB, Peele JL, Carelli RM and Wightman RM Sources contributing to the average extracellular concentration of dopamine in the nucleus accumbens. J Neurochem 121, 252–62 (2012).doi:10.1111/j.1471-4159.2012.07677.x [PubMed: 22296263]

29. Yapo C, Nair AG, Clement L, Castro LR, Hellgren Kotaleski J and Vincent P Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons. J Physiol (2017).doi:10.1113/JP274475

30. Freed CR and Yamamoto BK Regional brain dopamine metabolism: a marker for the speed, direction, and posture of moving animals. Science 229, 62–65 (1985). [PubMed: 4012312]

31. Niv Y, Daw ND, Joel D and Dayan P Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl) 191, 507–20 (2007).doi:10.1007/s00213-006-0502-4 [PubMed: 17031711]

32. Strecker RE, Steinfels GF and Jacobs BL Dopaminergic unit activity in freely moving cats: lack of relationship to feeding, satiety, and glucose injections. Brain Res 260, 317–21 (1983). [PubMed: 6831204]

33. Cohen JY, Amoroso MW and Uchida N Serotonergic neurons signal reward and punishment on multiple timescales. Elife 4, (2015).doi:10.7554/eLife.06346

34. Floresco SB, West AR, Ash B, Moore H and Grace AA Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. Nat Neurosci 6, 968–73 (2003).doi:10.1038/nn1103 [PubMed: 12897785]

35. Grace AA Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. Nature Reviews Neuroscience 17, 524 (2016).doi:10.1038/nrn.2016.57 [PubMed: 27256556]

36. Phillips PE, Stuber GD, Heien ML, Wightman RM and Carelli RM Subsecond dopamine release promotes cocaine seeking. Nature 422, 614–8 (2003).doi:10.1038/nature01476 [PubMed: 12687000]

37. Wassum KM, Ostlund SB and Maidment NT Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. Biol Psychiatry 71, 846–54 (2012).doi:10.1016/j.biopsych.2011.12.019 [PubMed: 22305286]

38. Howe MW, Tierney PL, Sandberg SG, Phillips PE and Graybiel AM Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature 500, 575–9 (2013).doi:10.1038/nature12475 [PubMed: 23913271]

39. Satoh T, Nakai S, Sato T and Kimura M Correlated coding of motivation and outcome of decision by dopamine neurons. J Neurosci 23, 9913–23 (2003). [PubMed: 14586021]

40. Howe MW and Dombeck DA Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature 535, 505–10 (2016).doi:10.1038/nature18942 [PubMed: 27398617]

41. Silva JAD, Tecuapetla F, Paixão V and Costa RM Dopamine neuron activity before action initiation gates and invigorates future movements. Nature 554, 244 (2018).doi:10.1038/nature25457 [PubMed: 29420469]

42. du Hoffmann J and Nicola SM Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens. J Neurosci 34, 14349–64 (2014).doi:10.1523/JNEUROSCI. 3492-14.2014 [PubMed: 25339748]

43. Hart AS, Rutledge RB, Glimcher PW and Phillips PE Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J Neurosci 34, 698–704 (2014).doi:10.1523/JNEUROSCI.2489-13.2014 [PubMed: 24431428]

44. Soares S, Atallah BV and Paton JJ Midbrain dopamine neurons control judgment of time. Science 354, 1273–1277 (2016).doi:10.1126/science.aah5234 [PubMed: 27940870]

45. Ikemoto S Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. Brain Res Rev 56, 27–78 (2007).doi:10.1016/ j.brainresrev.2007.05.004 [PubMed: 17574681]

46. Syed EC, Grima LL, Magill PJ, Bogacz R, Brown P and Walton ME Action initiation shapes mesolimbic dopamine encoding of future rewards. Nat Neurosci (2015).doi:10.1038/nn.4187

47. Floresco SB, Yang CR, Phillips AG and Blaha CD Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. Eur J Neurosci 10, 1241–51 (1998). [PubMed: 9749778]

48. Jones JL, Day JJ, Aragona BJ, Wheeler RA, Wightman RM and Carelli RM Basolateral amygdala modulates terminal dopamine release in the nucleus accumbens and conditioned responding. Biol Psychiatry 67, 737–44 (2010).doi:S0006–3223(09)01327–4 [pii] 10.1016/j.biopsych.2009.11.006 [PubMed: 20044074]

49. Cachope R, Mateo Y, Mathur BN, Irving J, Wang HL, Morales M, Lovinger DM and Cheer JF Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. Cell Rep 2, 33–41 (2012).doi:10.1016/j.celrep.2012.05.011 [PubMed: 22840394]

50. Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K and Cragg SJ Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. Neuron 75, 58–64 (2012).doi: 10.1016/j.neuron.2012.04.038 [PubMed: 22794260]

51. Grace AA Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41, 1–24 (1991). [PubMed: 1676137]

52. Moyer JT, Wolf JA and Finkel LH Effects of dopaminergic modulation on the integrative properties of the ventral striatal medium spiny neuron. J Neurophysiol 98, 3731–48 (2007). [PubMed: 17913980]

53. J drzejewska-Szmek J, Damodaran S, Dorman DB and Blackwell KT Calcium dynamics predict direction of synaptic plasticity in striatal spiny projection neurons. Eur J Neurosci 45, 1044–1056 (2017).doi:10.1111/ejn.13287 [PubMed: 27233469]

54. Morris G, Arkadir D, Nevet A, Vaadia E and Bergman H Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. Neuron 43, 133–43 (2004). [PubMed: 15233923]

55. Brown MT, Tan KR, O'Connor EC, Nikonenko I, Muller D and Lüscher C Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. Nature (2012).doi:10.1038/nature11657

56. Yamanaka K, Hori Y, Minamimoto T, Yamada H, Matsumoto N, Enomoto K, Aosaki T, Graybiel AM and Kimura M Roles of centromedian parafascicular nuclei of thalamus and cholinergic interneurons in the dorsal striatum in associative learning of environmental events. J Neural Transm (Vienna) (2017).doi:10.1007/s00702-017-1713-z

57. Shen W, Plotkin JL, Francardo V, Ko WK, Xie Z, Li Q, Fieblinger T, Wess J, Neubig RR, Lindsley CW, Conn PJ, Greengard P, Bezard E, Cenci MA and Surmeier DJ M4 Muscarinic Receptor Signaling Ameliorates Striatal Plasticity Deficits in Models of L-DOPA-Induced Dyskinesia. Neuron 88, 762–73 (2015).doi:10.1016/j.neuron.2015.10.039 [PubMed: 26590347]

58. Nair AG, Gutierrez-Arenas O, Eriksson O, Vincent P and Hellgren Kotaleski J Sensing Positive versus Negative Reward Signals through Adenylyl Cyclase-Coupled GPCRs in Direct and Indirect Pathway Striatal Medium Spiny Neurons. J Neurosci 35, 14017–30 (2015).doi:10.1523/JNEUROSCI.0730-15.2015 [PubMed: 26468202]

59. Stocco A Acetylcholine-based entropy in response selection: a model of how striatal interneurons modulate exploration, exploitation, and response variability in decision-making. Frontiers in neuroscience 6, (2012).

60. Franklin NT and Frank MJ A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. eLife 4, (2015).doi:10.7554/eLife.12029

61. Nougaret S and Ravel S Modulation of Tonically Active Neurons of the Monkey Striatum by Events Carrying Different Force and Reward Information. J Neurosci 35, 15214–26 (2015).doi:10.1523/JNEUROSCI.0039-15.2015 [PubMed: 26558790]

62. Schultz W Predictive reward signal of dopamine neurons. J Neurophysiol 80, 1–27 (1998). [PubMed: 9658025]

63. Lammel S, Hetzel A, Häckel O, Jones I, Liss B and Roeper J Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. Neuron 57, 760–73 (2008).doi:10.1016/j.neuron.2008.01.022 [PubMed: 18341995]

64. Poulin JF, Zou J, Drouin-Ouellet J, Kim KY, Cicchetti F and Awatramani RB Defining midbrain dopaminergic neuron diversity by single-cell gene expression profiling. Cell Rep 9, 930–43 (2014).doi:10.1016/j.celrep.2014.10.008 [PubMed: 25437550]

65. Morales M and Margolis EB Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. Nat Rev Neurosci 18, 73–85 (2017).doi:10.1038/nrn.2016.165 [PubMed: 28053327]

66. Matsumoto M and Hikosaka O Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459, 837–41 (2009).doi:nature08028 [pii] 10.1038/nature08028 [PubMed: 19448610]

67. Pasquereau B and Turner RS Dopamine neurons encode errors in predicting movement trigger occurrence. Journal of Neurophysiology 113, 1110–1123 (2014).doi:10.1152/jn.00401.2014 [PubMed: 25411459]

68. Redgrave P, Prescott TJ and Gurney K Is the short-latency dopamine response too short to signal reward error? Trends Neurosci 22, 146–51 (1999). [PubMed: 10203849]

69. Bromberg-Martin ES, Matsumoto M and Hikosaka O Dopamine in motivational control: rewarding, aversive, and alerting. Neuron 68, 815–34 (2010).doi:10.1016/j.neuron.2010.11.022 [PubMed: 21144997]

70. Dodson PD, Dreyer JK, Jennings KA, Syed EC, Wade-Martins R, Cragg SJ, Bolam JP and Magill PJ Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. Proc Natl Acad Sci U S A 113, E2180–8 (2016).doi:10.1073/pnas.1515941113 [PubMed: 27001837]

71. Lerner TN, Shilyansky C, Davidson TJ, Evans KE, Beier KT, Zalocusky KA, Crow AK, Malenka RC, Luo L, Tomer R and Deisseroth K Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. Cell 162, 635–47 (2015).doi:10.1016/j.cell.2015.07.014 [PubMed: 26232229]

72. Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND and Witten IB Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat Neurosci (2016).doi:10.1038/nn.4287

73. Kim CK, Yang SJ, Pichamoorthy N, Young NP, Kauvar I, Jennings JH, Lerner TN, Berndt A, Lee SY, Ramakrishnan C, Davidson TJ, Inoue M, Bito H and Deisseroth K Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. Nature Methods 13, 325–328 (2016).doi:10.1038/nmeth.3770 [PubMed: 26878381]

74. Menegas W, Babayan BM, Uchida N and Watabe-Uchida M Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. Elife 6, (2017).doi:10.7554/eLife.21886

75. Brown HD, McCutcheon JE, Cone JJ, Ragozzino ME and Roitman MF Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. The European journal of neuroscience 34, 1997–2006 (2011).doi:10.1111/j.1460-9568.2011.07914.x [PubMed: 22122410]

76. Knutson B and Greer SM Anticipatory affect: neural correlates and consequences for choice. Philos Trans R Soc Lond B Biol Sci 363, 3771–86 (2008).doi:10.1098/rstb.2008.0155 [PubMed: 18829428]

77. Bartra O, McGuire JT and Kable JW The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76, 412–27 (2013).doi:10.1016/j.neuroimage.2013.02.063 [PubMed: 23507394]

78. Ferenczi EA, Zalocusky KA, Liston C, Grosenick L, Warden MR, Amatya D, Katovich K, Mehta H, Patenaude B, Ramakrishnan C, Kalanithi P, Etkin A, Knutson B, Glover GH and Deisseroth K Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. Science 351, aac9698 (2016).doi:10.1126/science.aac9698 [PubMed: 26722001]

79. Bertran-Gonzalez J, Bosch C, Maroteaux M, Matamales M, Hervé D, Valjent E and Girault JA Opposing patterns of signaling activation in dopamine D1 and D2 receptor-expressing striatal neurons in response to cocaine and haloperidol. J Neurosci 28, 5671–85 (2008).doi:10.1523/JNEUROSCI.1039-08.2008 [PubMed: 18509028]

80. Redgrave P, Prescott TJ and Gurney K The basal ganglia: a vertebrate solution to the selection problem? Neuroscience 89, 1009–23 (1999). [PubMed: 10362291]

81. Beeler JA, Frazier CR and Zhuang X Putting desire on a budget: dopamine and energy expenditure, reconciling reward and resources. Front Integr Neurosci 6, 49 (2012).doi:10.3389/fnint.2012.00049 [PubMed: 22833718]

82. Anderson BA, Kuwabara H, Wong DF, Gean EG, Rahmim A, Braši JR, George N, Frolov B, Courtney SM and Yantis S The Role of Dopamine in Value-Based Attentional Orienting. Curr Biol 26, 550–5 (2016).doi:10.1016/j.cub.2015.12.062 [PubMed: 26877079]

83. Chatham CH, Frank MJ and Badre D Corticostriatal output gating during selection from working memory. Neuron 81, 930–42 (2014).doi:10.1016/j.neuron.2014.01.002 [PubMed: 24559680]

84. Shenhav A, Botvinick MM and Cohen JD The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79, 217–40 (2013).doi:10.1016/j.neuron.2013.07.007 [PubMed: 23889930]

85. Aarts E, Roelofs A, Franke B, Rijpkema M, Fernández G, Helmich RC and Cools R Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence

from genetic imaging. Neuropsychopharmacology 35, 1943–51 (2010).doi:10.1038/npp.2010.68 [PubMed: 20463658]

86. Westbrook A and Braver TS Dopamine Does Double Duty in Motivating Cognitive Effort. Neuron 89, 695–710 (2016).doi:10.1016/j.neuron.2015.12.029 [PubMed: 26889810]

87. Manohar SG, Chong TT, Apps MA, Batla A, Stamelou M, Jarman PR, Bhatia KP and Husain M Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. Curr Biol 25, 1707–16 (2015).doi:10.1016/j.cub.2015.05.038 [PubMed: 26096975]

88. Wunderlich K, Smittenaar P and Dolan RJ Dopamine Enhances Model-Based over Model-Free Choice Behavior. Neuron 75, 418–24 (2012).doi:10.1016/j.neuron.2012.03.042 [PubMed: 22884326]

89. Nicola SM The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. J Neurosci 30, 16585–600 (2010).doi:10.1523/JNEUROSCI.3958-10.2010 [PubMed: 21147998]

90. Eban-Rothschild A, Rothschild G, Giardino WJ, Jones JR and de Lecea L VTA dopaminergic neurons regulate ethologically relevant sleep-wake behaviors. Nat Neurosci (2016).doi:10.1038/nn. 4377

91. Haber SN, Fudge JL and McFarland NR Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci 20, 2369–82 (2000). [PubMed: 10704511]

92. Reddi BAJ and Carpenter RHS The influence of urgency on decision time. Nature neuroscience 3, 827 (2000). [PubMed: 10903577]

93. Thura D and Cisek P The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment. Neuron (2017).doi:10.1016/j.neuron.2017.07.039

94. Turner RS and Desmurget M Basal ganglia contributions to motor control: a vigorous tutor. Curr Opin Neurobiol 20, 704–16 (2010).doi:10.1016/j.conb.2010.08.022 [PubMed: 20850966]

95. Hikosaka O, Nakamura K and Nakahara H Basal ganglia orient eyes to reward. J Neurophysiol 95, 567–84 (2006).doi:10.1152/jn.00458.2005 [PubMed: 16424448]

96. Kelly PH and Moore KE Mesolimbic dopaminergic neurones in the rotational model of nigrostriatal function. Nature 263, 695–6 (1976). [PubMed: 980114]

97. Cousins MS, Atherton A, Turner L and Salamone JD Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. Behav Brain Res 74, 189–97. (1996). [PubMed: 8851929]

98. Redish AD Vicarious trial and error. Nat Rev Neurosci 17, 147–59 (2016).doi:10.1038/nrn.2015.30 [PubMed: 26891625]

99. Rabinovich MI, Huerta R, Varona P and Afraimovich VS Transient cognitive dynamics, metastability, and decision making. PLoS Comput Biol 4, e1000072 (2008).doi:10.1371/journal.pcbi.1000072 [PubMed: 18452000]

100. Merchant H, Harrington DL and Meck WH Neural basis of the perception and estimation of time. Annu Rev Neurosci 36, 313–36 (2013).doi:10.1146/annurev-neuro-062012-170349 [PubMed: 23725000]
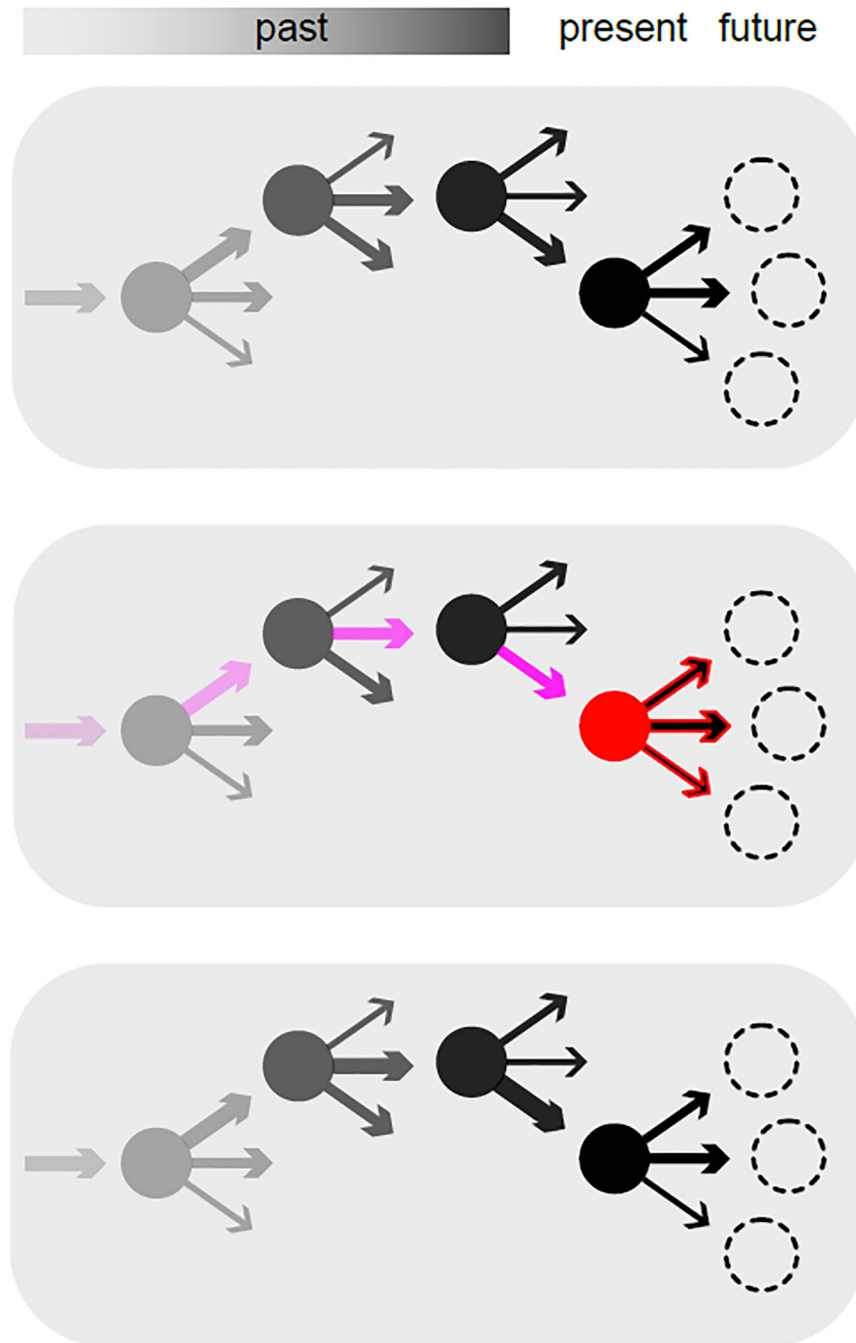
FIG. 1. Dopamine: updating the past, invigorating the present.
Top, Circles with arrows represent states and the potential actions from those states. Arrow widths indicate learned values of performing each action. As states/actions fade into the past, they are progressively less eligible for reinforcement. *Middle*, a burst of dopamine occurs. The result is invigoration of actions available from the current state (red), and plasticity of the value representations for recently performed actions (purple). *Bottom*, as the result of plasticity, the next time these states are encountered their associated values have increased (arrow widths). Through repeated experience reinforcement learning can "carve a

groove" through state space, making certain trajectories increasingly more likely. In addition to this learning role, the invigorating, performance role of dopamine seems to speed up the flow along previously-learned trajectories.
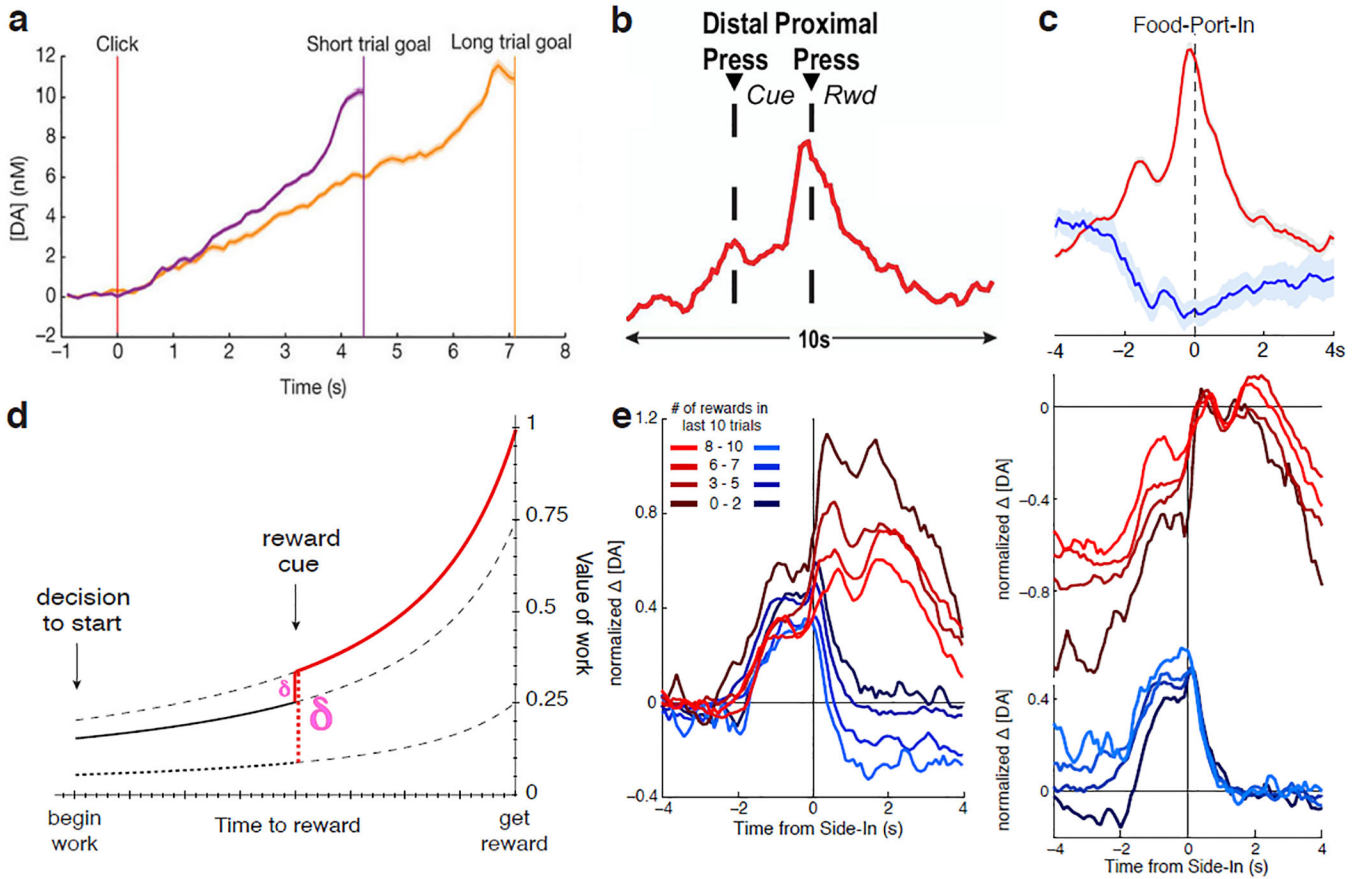
FIG. 2. Fast dopamine fluctuations signal dynamically-evolving reward expectations.

a-c) Mesolimbic dopamine release rapidly increases as rats get closer to anticipated rewards. d) Value, defined as temporally-discounted estimates of future reward, increases as reward gets closer. Cues indicating that reward is larger, closer, or more certain than previously expected cause jumps in value. These jumps from one moment to the next are temporal-difference RPEs. e) Subtracting away "baselines" can confound value and RPE signals. Left, dopamine aligned to reward-predictive cue(at time zero), with conventional baseline subtraction, seems to show that dopamine jumps to higher levels when reward is less expected(brown), resembling an RPE signal. Right, an alternative presentation of the same data, equating dopamine levels *after* the cue, would show instead that precue dopamine levels depend on reward expectation(value). Additional analyses determined that the right-side presentation is closer to the truth(see details in ref. 15). Panel a reproduced, with permission, from ref 38, Macmillan Publishers Limited….; panel b reproduced, with permission, from ref. 37, Elsevier; panels c-e reproduced, with permission, from ref 15, Macmillan Publishers Limited