Abundance compensates kinetics: Similar effect of dopamine signals on D1 and D2 receptor populations

Lars Hunger[1*], **Arvind Kumar**[2], **Robert Schmidt**[1]

***For correspondence:**
lhunger1@sheffield.ac.uk;
Lars.Hunger.314@gmail.com

[1]Department of Psychology, University of Sheffield; [2]Department of Computational Science and Technology, KTH Stockholm

Abstract The properties of different dopamine receptors constrain the function of dopamine signals in the striatum of the basal ganglia. Still, dopamine receptor kinetics are often neglected in considerations of the temporal dynamics of dopamine signalling. Here we develop a neurochemical model of dopamine receptor binding taking into account slow receptor kinetics. Contrary to current views, in our model D1 and D2 dopamine receptor populations react very similarly to dopamine signals independent of their timescale and integrate them over minutes. Furthermore, our model explains why ramping dopamine concentrations, observed experimentally, are an effective signal for increasing the occupancy of dopamine receptors.

Introduction

The neuromodulator dopamine (DA) has complex effects on the activity of striatal neurons by changing their excitability (***Day et al., 2008***) and strength of synaptic inputs (***Reynolds et al., 2001***) in the context of motor control (***Syed et al., 2016***), action-selection (***Redgrave et al., 2010***), reinforcement learning (***Schultz, 2007***), and addiction (***Everitt and Robbins, 2005***). Striatal DA concentration ([DA]) may change over multiple timescales (***Schultz, 2007***). Fast, abrupt increases in [DA] lasting for $\approx 1-3s$ result from phasic bursts in DA neurons (***Roitman et al., 2004***), which signal reward-related information (***Schultz, 2007***; ***Grace et al., 2007***). Slightly slower [DA] ramps occur when rats approach a goal location (***Howe et al., 2013***) or perform a reinforcement learning task (***Hamid et al., 2016***). Finally, slow tonic spontaneous firing of DA neurons controls the baseline [DA] and may change on a timescale of minutes or longer (***Grace et al., 2007***). However, whether fast and slow changes in [DA] actually represent distinct signalling modes, e.g. for learning and motivation (***Niv et al., 2007***), has recently been challenged (***Berke, 2018***). Furthermore, DA acts on two different main receptor types, D1 and D2, adding another layer of complexity to its signalling.

Based on different DA affinities of D1 and D2 receptors (D1R and D2R), it is often assumed that striatal medium spiny neurons (MSN) respond differently to tonic and phasic DA changes, depending on which DA receptor they express predominantly (***Dreyer et al., 2010***; ***Surmeier et al., 2007***; ***Grace et al., 2007***; ***Schultz, 2007***; ***Frank and O'Reilly,***

2006). According to this "affinity-based" model the low affinity D1Rs (i.e. high dissociation constant $K_D^{D1} = 1.6\mu M$; *Richfield et al., 1989*) cannot detect tonic changes in [DA] because the fraction of occupied D1Rs is small ($\approx 1\%$) at baseline [DA] ($20nM$), see Methods) and does not change much during tonic, low amplitude [DA] changes. However, D1Rs seem well suited to detect phasic, high amplitude [DA] increases because they saturate at very high [DA]. By contrast, D2Rs have a high affinity (i.e. low dissociation constant $K_D^{D1} = 25nM$; *Richfield et al., 1989*) leading to $\approx 40\%$ of D2Rs being occupied at baseline [DA] ($20nM$). Due to their high affinity, D2Rs can detect low amplitude, tonic increases/decreases in [DA]. However, as D2Rs saturate at a relatively low $[DA] > 2 \cdot K_D^{D2}$, they seem unable to detect high amplitude, phasic increases in [DA]. This suggests that D1 and D2 type MSNs differentially encode phasic and tonic changes in [DA] solely because of the different affinities of D1Rs and D2Rs (*Schultz, 2007*). However, this view is incompatible with recent findings that D2R expressing MSNs can detect phasic changes in [DA] (*Yapo et al., 2017*; *Marcott et al., 2014*).

The affinity-based model assumes that the reaction equilibrium is reached instantaneously, whereby the receptor binding affinity can be used to approximate the fraction of receptors bound to DA. However, this assumption holds only if the receptor kinetics are fast with respect to the timescale of the DA signal, which is typically not the case. For instance, D1Rs and D2Rs unbind from DA with a half-life time of $t_{1/2} \approx 80s$ (*Burt et al., 1976*; *Sano et al., 1979*; *Maeno, 1982*; *Nishikori et al., 1980*), much longer than phasic signals of a few seconds (*Robinson et al., 2001*; *Schultz, 2007*; *Hamid et al., 2016*). Moreover, the fraction of bound receptors might be a misleading measure for the effect of DA signals, since the abundances of D1R and D2R are quite different (see below). Therefore, we developed a model of receptor binding based on the kinetics and abundances of D1Rs and D2Rs to re-evaluate current views on DA signalling in the striatum.

Results and Discussion

To provide a realistic description of receptor kinetics, the binding and unbinding rates that determine the receptor affinity are required. The available experimental measurements indicate that the different D1R and D2R affinities are largely due to different binding rates, while their unbinding rates are similar (*Burt et al., 1976*; *Sano et al., 1979*; *Maeno, 1982*; *Richfield et al., 1989*). We incorporated these measurements into our slow kinetics model (see Methods) and investigated the model in a variety of scenarios mimicking DA signals on different timescales.

Firstly, to examine our model at baseline [DA], we investigated receptor binding for a range of affinities (**Fig. 1a**), reflecting the range of measured values in different experimental studies (*Neve and Neve, 1997*). We report the resulting receptor occupancy in terms of the concentration of D1Rs and D2Rs bound to DA (denoted as $[D1 - DA]$ and $[D2 - DA]$, respectively). Due to the low affinity of D1Rs, slow changes in [DA] only lead to small changes in the fraction of bound D1 receptors. However, there are overall more D1Rs than D2Rs (*Richfield et al., 1989*), and $\approx 80\%$ of D2Rs are retained in the endoplasmatic reticulum (*Prou et al., 2001*). Therefore, the concentration of D1Rs in the membrane available to extracellular DA is a lot higher than the concentration of D2Rs (e.g. 20 times more in the nucleus accumbens; *Nishikori et al., 1980*; Methods). Thus, in our simulation, the actual concentration of bound D1Rs ($[D1 - DA] \approx 20nM$) was, at DA baseline, much closer to the concentration of bound D2Rs ($[D2 - DA] \approx 35nM$) than

suggested by the different D1 and D2 affinities alone. We further confirmed that this was not due to a specific choice of the dissociation constants in the model, as $[D1-DA]$ and $[D2-DA]$ remained similar over the range of experimentally measured D1R and D2R affinities (***Neve and Neve, 1997***) (**Fig. 1a**). This suggests that $[D1-DA]$ is at most twice as high as $[D2-DA]$ instead of 40 times higher as suggested by the difference in fraction of bound receptors. Therefore, $[D1-DA]$ and $[D2-DA]$ might be better indicators for the signal transmitted to MSNs as the fraction of bound receptors neglects the different receptor type abundances.

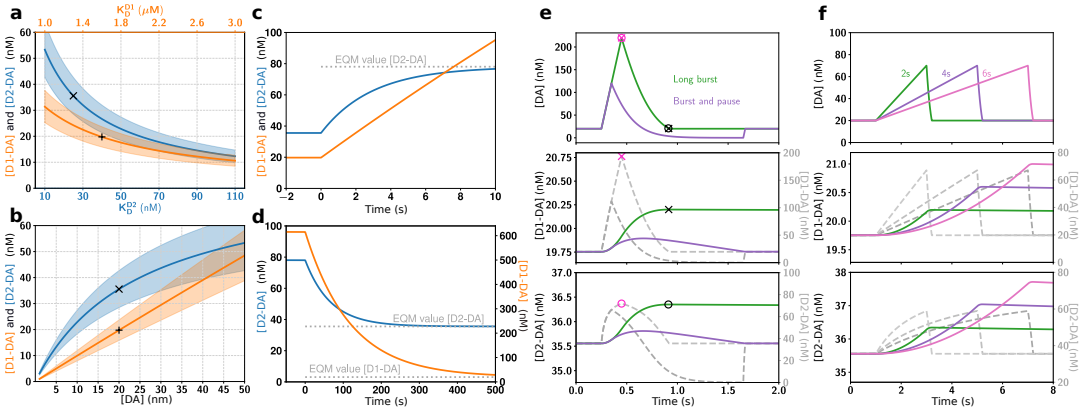


Figure 1. Impact of slow kinetics on D1R and D2R binding. **(a, b)** Equilibrium values of absolute concentration of receptors bound to DA vary as a function of receptor affinities (a) and baseline [DA] (b), but are overall similar for D1 and D2 receptors. In (a) baseline [DA] was fixed at 20 nM, and in (b) $K_D^{D1} = 1.6\mu M$ and $K_D^{D2} = 25nM$. 'x' and +' indicate the model default parameters. Coloured bands mark the range of values for up to $\pm 20\%$ different receptor abundances. **(c)** For a large step up from $[DA] = 20nM$ to $[DA] = 1\mu M$, and **(d)** a step down from $[DA] = 1\mu M$ to $[DA] = 20nM$, D1 and D2 receptor occupancy approached their new equilibrium (EQM, grey dotted lines) only slowly (i.e. over seconds to minutes). **(e, f)** Effect of different phasic DA signals (top panels) is very different in the slow kinetics model (coloured traces in middle and bottom panels; left scales) compared to the instant kinetics model (dashed grey traces, right scales). The timing of the maximum receptor occupancy ('x' and 'o' for D1 and D2, respectively) coincides for instant kinetics (purple symbols) with the [DA] peak (combined x and o in top panel), while for slow kinetics (black symbols) it coincides with the offset of the [DA] signal instead (combined x and o in top panel).

Next, we investigated the effect of slow [DA] changes (***Grace, 1995***; ***Schultz, 1998***; ***Floresco et al., 2003***) by exposing our model to changes in the [DA] baseline. For signalling timescales that are long with respect to the half-life time of the receptors ($t_{slow} >> t_{1/2} \approx 80s$), we used the dissociation constant to calculate the steady state receptor occupancy. We found that for slow changes to a range of [DA] baselines, $[D1-DA]$ and $[D2-DA]$ were also similar (**Fig. 1b**). Thus, we conclude that D1R and D2R occupancy reacts similarly to slow, low amplitude [DA] changes because of the different abundances of D1 and D2 receptors. This is contrary to instant kinetics models, suggesting that D2Rs are better suited to encode slow or tonic changes in [DA].

To study the impact of faster [DA] signals, we measured the step response of the model to a [DA] change from 20nM to $1\mu M$. This is quite a large change compared to phasic DA signals in vivo (***Robinson et al., 2001***; ***Cheer et al., 2007***; ***Hamid et al., 2016***), which we choose to illustrate that our results are not just due to a small amplitude DA signal. We found that binding to both receptor subtypes increased very slowly. Even for

the high affinity D2Rs it took more than $5s$ to reach their new equilibrium (**Fig. 1c**). Thus, unlike the instant kinetics model, our model suggests that the D2Rs will not saturate for single reward events, which last overall for up to $\approx 3s$. Note that the non-saturation is independent of the abundance of the receptors and is only determined by the kinetics of the receptors (see Methods). Due to their slow unbinding, D1Rs and D2Rs also took a long time to return to baseline receptor occupancy after a step down from $[DA] = 1\mu M$ to $[DA] = 20nM$ (**Fig. 1d**). Thus, we conclude that with slow kinetics of receptor binding both D1Rs and D2Rs can detect single phasic DA signals and that both remain occupied long after the [DA] has returned to baseline.

Next, we investigated $[D1 - DA]$ and $[D2 - DA]$ for a phasic DA increase (mimicking reward responses; ***Robinson et al., 2001***; ***Cheer et al., 2007***), a phasic DA increase followed by a decrease (mimicking responses to non-reward, salient stimuli; ***Schultz, 2016***), and a prolonged DA ramp (mimicking goal approach; ***Howe et al., 2013***; ***Hamid et al., 2016***). In the instant kinetics model the D1Rs mirrored the [DA] time course, since even at $[DA] = 200nM$ they are far from saturation, whereas the D2Rs showed saturation effects as soon as $[DA] > 2 \cdot K_D^{D2}$, leading to differing D1 and D2 time courses (**Fig. 1e, f**). Importantly, in our model with slow kinetics, the time courses of $[D1 - DA]$ and $[D2 - DA]$ were similar for each of the three types of phasic DA signals.

While in our model we assumed slow kinetics based on neurochemical estimates of wildtype DA receptors (***Burt et al., 1976***; ***Sano et al., 1979***; ***Maeno, 1982***), recent genetically-modified DA receptors, used to probe [DA] changes, have apparent fast kinetics (***Sun et al., 2018***; ***Patriarchi et al., 2018***). Although their kinetics strongly changed between receptor variants and may not reflect the kinetics of the wildtype receptor, we examined our model also in the context of faster DA kinetics and found that the similarity between $[D1 - DA]$ and $[D2 - DA]$ can be observed even if the actual kinetics were a 100 times faster than assumed in our model (**Supp. Fig. 1**). Therefore, our results do not depend on the exact kinetics parameters or potential temperature effects, as long as the parameter changes are roughly similar for D1 and D2 receptors. Furthermore, taking into account different affinity states for D1Rs and D2Rs (***Richfield et al., 1989***), preserved the similarity of time courses of D1R and D2R occupancy (**Supp. Fig. 7**). Finally, pauses in the DA firing following aversive stimuli (***Schultz, 2007***) that lead to reductions in [DA] (***Roitman et al., 2008***), also have a similar effect on D1R and D2R occupancy (**Supp. Fig. 4e**).

Another striking effect of incorporating receptor kinetics was that a phasic increase in [DA] kept the receptors occupied for a long time (**Fig. 1e**). However, when a phasic increase was followed by a decrease, $[D1 - DA]$ and $[D2 - DA]$ quickly returned to baseline. This indicates that burst-pause firing patterns observed in DA cells for aversive or salient non-rewarding signals (***Schultz, 2016***) can be distinguished from pure burst firing patterns (which only lead to a phasic increase in [DA]) on the level of the MSN DA receptor occupancy. This supports the view that the fast component of the DA firing patterns (***Schultz, 2016***) is a salience response, and points to the intriguing possibility that the pause following the burst can, at least partly, revoke the receptor-ligand binding induced by the burst (see also **Supp. Fig.2**). This effect even persists in a sequence of burst and burst-pause events (**Supp. Fig. 5**). Thereby, the burst-pause firing pattern of DA neurons could effectively signal a reward false-alarm.

The similarity of $[D1 - DA]$ and $[D2 - DA]$ responses to both slow and fast [DA] changes indicates that the different DA receptors respond similarly independent of the timescale

of [DA] changes. To understand why the D1Rs and D2Rs respond similarly, we considered the relevant model parameters in more detail. The binding rate constants of D1Rs and D2Rs differ by a factor of ≈ 60 ($k_{on}^{D1} = 0.0003125 nm^{-1}min^{-1}$ and $k_{on}^{D2} = 0.02 nm^{-1}min^{-1}$; ***Burt et al., 1976***; ***Sano et al., 1979***; ***Maeno, 1982***; Methods), suggesting faster D2Rs. However, experimental data suggests that there are ≈ 40 fold more unoccupied D1 receptors ($\approx 1600nM$) than unoccupied D2 receptors ($\approx 40nM$) on MSN membranes in the extracellular space of the rat striatum (***Nishikori et al., 1980***). Therefore, the absolute binding rate $\frac{d[DX-DA]}{dt}^{+} = k_{on} \cdot [DA] \cdot [DX]$ differs only by a factor of ≈ 1.5 between the D1Rs and D2Rs. That is, the difference in the kinetics of D1Rs and D2Rs is compensated by the different receptor numbers, resulting in nearly indistinguishable aggregate kinetics (**Fig. 1e, f**). This is consistent with recent experimental findings that D2R expressing MSNs can detect phasic [DA] signals (***Yapo et al., 2017***; ***Marcott et al., 2014***).

Incorporating the slow kinetics in the model is crucial for functional considerations of the DA system. Currently, following the instant kinetics model, the amplitude of a DA signal (i.e. peak [DA]) is often considered as a key signal e.g. in the context of reward magnitude or probability (***Hamid et al., 2016***; ***Tobler et al., 2005***; ***Morris et al., 2004***). However, as DA unbinds slowly (over tens of seconds; **Fig. 1d**) and the binding rate changes approximately linearly with [DA], the amount of receptor occupancy primarily depended on the area under the curve of the [DA] signal (**Supp. Fig. 3**). Therefore, DA ramps, even with a relatively small amplitude (**Fig. 1f** and **Supp. Fig. 4**), were very effective in increasing DA receptor binding. In contrast, for locally very high [DA] (e.g. at corticostriatal synapses during phasic DA cell activity; ***Grace et al., 2007***) the high concentration gradient would only lead to a very short duration of this local DA peak and thereby make it less effective in occupying DA receptors.

The dynamics introduced by the slow kinetics had further effects on the timecourse of DA signalling. With instant kinetics the maximum receptor occupancy was reached at the peak [DA] (**Fig. 1e, f**). By contrast, for slow kinetics the maximum receptor occupancy was reached when [DA] returned to its baseline (**Fig. 1e**) because as long as [DA] was higher than the equilibrium value of [D1-DA] and [D2-DA], more receptors continued to become occupied. Therefore for all DA signals, the maximum receptor occupancy was reached towards the end of the pulse (**Fig. 1e, f** and **Supp. Fig. 4**).

Another effect of the slow kinetics was that DA receptors remained occupied long after the DA pulse is over (**Fig. 1e, f**). This allowed the integration of DA pulses over minutes (**Fig. 2a, b** and **Supp. Fig.5**). We investigated potential functional consequences of this integration by exposing the model to a sequence of trials modeling a simple behavioural experiment with stochastic rewards (**see Methods**). We found that both D1R and D2R occupancy coded for reward probability (**Fig. 2** and **Supp. Fig. 6**), consistent with functional roles of DA signalling in motivation. However, this does not preclude potential DA roles on shorter time scales, such as the invigoration of movements (***Roesch et al., 2009***) or fast updates of state value (***Hamid et al., 2016***), as a sensitive readout mechanism could also detect small increases in [D1-DA] and [D2-DA] (***Lamb and Pugh Jr, 1992***).

Overall, our slow kinetics model of DA receptor binding casts doubt on several long-held views on DA signalling. Our model indicates that both D1R and D2R systems can detect [DA] changes, independent of the timescale, equally well. Although, D1Rs and D2Rs have opposing effects on the excitability (***Flores-Barrera et al., 2011***) and strength

of cortico-striatal synapse of D1 and D2 type MSNs (***Centonze et al., 2001***), we challenge the current view that differences in receptor affinity introduce additional asymmetries in D1 and D2 signalling. Instead of listening to different components of the DA signal, D1 and D2 MSNs seem to respond to the same DA input, increasing the differential effect on firing rate response of D1 and D2 MSNs.

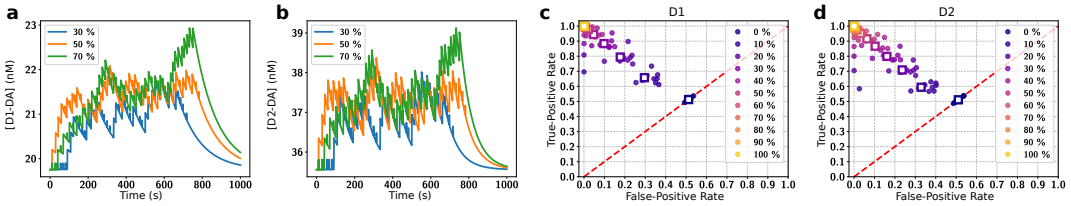


Figure 2. Integration of DA signals over minutes in a simulation of a behavioural task. **(a, b)** Timecourse of D1 (a) and D2 (b) receptor occupancy for sequences of 50 trials with a reward probability, as indicated, in each trial. Rewarded trials were modelled with a long DA burst, non-rewarded trials with a burst-pause DA timecourse. **(c, d)** Decoding accuracy of the difference in reward probability based on the D1 (c) and D2 (d) receptor occupancy by a simple classifier. Each data point indicates the decoding accuracy from a simulation scenario with the difference in reward probability indicated by the colour. Single dots correspond to simulations with different absolute reward probabilities. The colour indicates the difference in reward probability (e.g. a 10% difference in purple occurs for 80% vs. 90%, 70% vs. 80%, etc.), and the squares denote the corresponding averages. Red line indicates chance level performance, and a perfect classifier would be at 1.0 true and 0.0 false positive rate. Note that the classification is similar for D1 and D2 receptors, yielding near perfect classification already at 40% reward difference.

Methods and Materials

The models were implemented in Python. The scripts used to generate the data and figures can be accessed here: https://bitbucket.org/Narur/abundance_kinetics/src/.

Kinetics model

In the instant kinetics model the fraction of occupied D1 and D2 receptors (f_{D1} and f_{D2}) are calculated directly from the concentration of free DA in the extracellular space, $[DA]$, and the dissociation constant K_D:

$$f = \frac{[DA]}{K_D + [DA]}. \tag{1}$$

However, the dissociation constant is an equilibrium constant, so it should only be used for calculating the receptor occupancy when the duration of the DA signal is longer than the time needed to reach the equilibrium. As this is typically not the case for phasic DA signals (see main text), we developed a model incorporating slow kinetics.

When DA and one of its receptors are both present in a solution they constantly bind and unbind. During the binding a receptor ligand complex (here called DA–D1 or DA–D2) is formed. We call the receptor ligand complex an occupied DA receptor. Note that although in the following part we provide the equations for D1 receptors, the same equations apply for D2 receptors (with different kinetic parameters). In a solution binding

occurs when receptor and ligand meet due to diffusion, with high enough energy and a suitable orientation, described as:

$$\mathrm{DA} + \mathrm{D1} \xrightarrow{k_{on}} \mathrm{DA{-}D1}. \tag{2}$$

Accordingly, unbinding of the complex is denoted as:

$$\mathrm{DA{-}D1} \xrightarrow{k_{off}} \mathrm{DA} + \mathrm{D1}. \tag{3}$$

The kinetics of this binding and unbinding, treated here as first-order reactions, are governed by the rate constants k_{on} and k_{off} that are specific for a receptor ligand pair and temperature dependent. Since both processes are happening simultaneously we can write this as:

$$\mathrm{DA} + \mathrm{D1} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{DA{-}D1}. \tag{4}$$

The rate at which the receptor is occupied depends on [DA], the concentration of free receptor [D1] and the binding rate constant k_{on}:

$$\frac{d[\mathrm{DA-D1}]^{+}}{dt} = k_{\mathrm{on}} \cdot [\mathrm{DA}] \cdot [\mathrm{D1}]. \tag{5}$$

The rate at which the receptor-ligand complex unbinds is given by concentration of the complex $[\mathrm{DA-D1}]$ and the unbinding rate constant k_{off}:

$$\frac{d[\mathrm{DA-D1}]^{-}}{dt} = -k_{\mathrm{off}} \cdot [\mathrm{DA-D1}]. \tag{6}$$

The equilibrium is reached when the binding and unbinding rates are equal, so by combining Eq. 5 and Eq. 6 we obtain:

$$k_{\mathrm{on}} \cdot [\mathrm{DA}] \cdot [\mathrm{D1}] = k_{\mathrm{off}} \cdot [\mathrm{DA-D1}]. \tag{7}$$

At the equilibrium the dissociation constant K_D is defined as:

$$K_D = \frac{[DA] \cdot [D1]}{[DA-D1]} = \frac{k_{off}}{k_{on}}. \tag{8}$$

When half of the receptors are occupied, i.e. $[DA{-}D1] = [D1]$, Eq. 8 simplifies to $K_D = [DA]$. So at equilibrium, K_D is the ligand concentration at which half of the receptors are occupied.

Importantly, for fast changes in [DA] (i.e. over seconds) it takes some time until the changed binding (Eq. 5) and unbinding rates (Eq. 6) are balanced, so the new equilibrium will not be reached instantly. The timescale in which equilibrium is reached can be estimated from the half-life time of the bound receptor. The half-life time assumes an exponential decay process as described in Eq. 6 and is the time required so that half of the currently bound receptors unbind. If $[DA] = 0$, and there is no more binding, the half life time of the receptors can be calculated from the off-rate by using $t_{1/2} = ln(2)/k_{off}$. Signal durations should be of the same order of magnitude (or longer) than the half-life time in order for the instant kinetics model to be applicable.

We calculated the time course of occupied receptor after an abrupt change in [DA] by integrating the rate equation, given by the sum of Eq. 5 and Eq. 6:

$$\frac{d[DA-D1]}{dt} = k_{on}[DA][D1] - k_{off}[DA-D1]. \tag{9}$$

To integrate Eq. 9 we substitute

$$[D1] = [D1^{tot}] - [DA - D1] \tag{10}$$

where $[D1^{tot}]$ is the total amount of D1 receptor (bound and unbound to DA) on the cell membranes available for binding to extracellular DA.

To model the effect of phasic changes in [DA] we choose the initial receptor occupancy $[DA - D1](t = 0) = [DA - D1]^0$ and the receptor occupancy for the new equilibrium at time infinity $[DA - D1](t = \infty) = [DA - D1]^\infty$ as the boundary conditions. With these boundary conditions we get an expression for the time evolution of the receptor occupancy under the assumption that binding to the receptor does not significantly change the free [DA]:

$$\begin{aligned}[DA - D1](t) = \\ ([DA - D1]^0 - [DA - D1]^\infty) \cdot e^{-(k_{on}[DA] + k_{off})t} \\ + [DA - D1]^\infty.\end{aligned} \tag{11}$$

For our slow kinetics model we solved Eq. 9 for each receptor type and arbitrary DA timecourses numerically employing a 4th order Runge Kutta solver with a 1 ms time resolution.

We did not take into account the change in [DA] caused by the binding and unbinding to the receptors since the rates at which DA is removed from the system by binding to the receptors is much slower than the rate of DA being removed from the system by uptake through DA transporters. For example the rate at which DA binds to the receptors is:

$$\begin{aligned}\frac{([DA - D1] + [DA - D2])}{dt} = \\ k_{on}^{D1}[DA][D1] + k_{on}^{D2}[DA][D2] \\ = -\frac{[DA]}{dt}.\end{aligned} \tag{12}$$

At a DA concentration of $[DA] = 1\mu M$ with a D1 and D2 occupancy of $[DA - D1] \approx 20.0nM$ and $[DA - D2] \approx 40nM$ (the equilibrium values for $[DA] = 20nM$) and $k_{on}^{D1} = 5.2 \cdot 10^{-6} nM^{-1}s^{-1}$, $k_{on}^{D2} = 3.3 \cdot 10^{-4} nM^{-1}s^{-1}$, $[D1] \approx 1600.0nM$, $[D2] \approx 40.0nM$ and $[DA] = 1\mu M$ the rate of DA removal through binding to the receptors is:

$$\frac{[DA]^{binding}}{dt} = -23.6nM/s. \tag{13}$$

However, the DA removal rate by Michaelis-Menten uptake through the DA transporters at this concentration would be:

$$\frac{[DA]^{uptake}}{dt} = V_{max}\frac{[DA]}{[DA] + K_m} \tag{14}$$

$$= -4.0\frac{\mu M}{s} \cdot \frac{1\mu M}{1\mu M + 0.21\mu M} \tag{15}$$

$$= -3.3\frac{\mu M}{s}. \tag{16}$$

Where V_{max} is the maximal uptake rate, and K_m the Michaelis-Menten constant describing the $[DA]$ concentration at which uptake is at half the maximum rate. As $\left|\frac{[DA]^{uptake}}{dt}\right| >>$

$\left|\frac{[DA]^{binding}}{dt}\right|$, the DA dynamics are dominated by the uptake process and not by binding to the receptors. Therefore, we neglected the receptor-ligand binding for the DA dynamics in our model. However, for faster DA receptors this effect would become more important.

Receptor parameters

An important model parameter is the total concentration of the D1 and D2 receptors on the membrane ($[D1]^{tot}$ and $[D2]^{tot}$) that can bind to DA in the extracellular space of the striatum. Our estimate of $[D1]^{tot}$ and $[D2]^{tot}$ is based on radioligand binding studies in the rostral striatum (***Richfield et al., 1989***, ***1987***). We use the following equation, in which X is a placeholder for the respective receptor type, to calculate these concentrations.

$$[DX]^{tot} = [DX]^{m} \cdot \frac{\epsilon \cdot f_{DX}^{membrane}}{\alpha \rho_{brain}} \tag{17}$$

The experimental measurements provide us with a number of receptors per unit of protein weight $[D1]^{m}$ and $[D2]^{m}$. To transform these measurements into molar concentrations for our simulations, we multiply by the protein content of the wet weight of the rat caudate nucleus ϵ, which is around 12% (***Banay-Schwartz et al., 1992***). This leaves us with the amount of protein per g of wet weight of the rat brain. Next we divide by the average density of a rat brain which is $\rho_{brain} = 1.05 g/ml$ (***DiResta et al., 1990***) to find the amount of receptors per unit of volume of the rat striatum. Finally, we divide by the volume fraction α, the fraction of the brain volume that is taken up by the extracellular space in the rat brain, to obtain the receptor concentration of the receptor in the extracellular medium. The procedure ends here for the D1 receptors since there is no evidence that D1 receptors are internalized in the baseline state (***Prou et al., 2001***). However, a large fraction of the D2 receptors is retained in the endoplasmatic reticulum of the neuron (***Prou et al., 2001***), reducing the amount of receptors that contribute to the concentration of receptors in the extracellular medium by $f^{membrane}$, the fraction of receptors protruding into the extracellular medium.

In addition to the receptor concentration, the kinetic constants of the receptors are key parameters in our slow kinetics model. In an equilibrium measurement in the canine caudate nucleus the dissociation constant of low affinity DA binding sites, corresponding to D1 receptors (***Maeno, 1982***), has been measured as $K_d = 1.6 \mu M$ (***Sano et al., 1979***). However, when calculating K_d (using Eq. 8) from the measured kinetic constants (***Sano et al., 1979***) the value is $K_d^{D1} = 2.6 \mu M$. To be more easily comparable to other simulation works (***Dreyer et al., 2010***) and direct measurements (***Richfield et al., 1989***; ***Sano et al., 1979***) we choose $K_d^{D1} = 1.6 \mu M$ in our simulations. For this purpose we modified both the $k_{on}^{D1} = 0.00025 min^{-1} nM^{-1}$ and $k_{off}^{D1} = 0.64 min^{-1}$ rate measured (***Sano et al., 1979***) by $\approx 25\%$, making $k_{on}^{D1} = 0.0003125 min^{-1} nM^{-1}$ slightly faster and $k_{off}^{D1} = 0.5 min^{-1}$ slightly slower, so that the resulting $K_d^{D1} = 1.6 \mu M$. The kinetic constants have been measured at $30\ C$ and are temperature dependent. In biological reactions a temperature change of $10\ C$ is usually associated with a change in reaction rate around a factor of 2-3 (***Reyes et al., 2008***). However, the conclusions of this paper do not change for an increase in reaction rates by a factor of $2-3$ (see **Supp. Fig. 1**). It should also be noted that the measurement of the commonly referenced K_d (***Richfield et al., 1989***) have been performed at room temperature.

The kinetic constants for the D2 receptors were obtained from measurements at $37^{o} C$

of high affinity DA binding sites (**Burt et al., 1976**), which correspond to the D2 receptor (**Maeno, 1982**). The values are $k_{on}^{D2} = 0.02 min^{-1} nM^{-1}$ and $k_{off}^{D2} = 0.5 min^{-1}$, which yields $K_D^{D2} = 25nM$, in line with the values measured in (**Richfield et al., 1989**). As the off-rate of the D1 and D2 receptors $k_{off}^{D1} = 0.64 min^{-1} nM^{-1}$ and $k_{off}^{D2} = 0.5 min^{-1}$ is quite similar, the difference in $K_D^{D2} = 25nM$ and $K_d^{D1} = 1.5 \mu M$ is largely due to differences in the on-rate of the receptors. This is important because the absolute rate of receptor occupancy depends linearly not only on the on-rate, but also on the receptor concentration (see Eq. 5), which means that a slower on-rate could be compensated for by a higher number of receptors.

Measured values		
Parameter		Source
$[D1]^m$ in pmol/mg protein	2840	(**Richfield et al., 1989**)
$[D2]^m$ in pmol/mg protein	696	(**Richfield et al., 1989**)
ϵ	0.12	(**Banay-Schwartz et al., 1992**)
α	0.2	(**Syková and Nicholson, 2008**)
ρ_{brain} in g/ml	1.05	(**DiResta et al., 1990**)
$f_{D1}^{membrane}$	1.0	(**Prou et al., 2001**)
$f_{D2}^{membrane}$	0.2	(**Prou et al., 2001**)
$k_{on}^{D1,orig}$ in $nm^{-1} min^{-1}$	0.00025	(**Sano et al., 1979**)
$k_{off}^{D1,orig}$ in min^{-1}	0.64	(**Sano et al., 1979**)
k_{on}^{D2} in $nm^{-1} min^{-1}$	0.02	(**Burt et al., 1976**)
k_{off}^{D2} in min^{-1}	0.5	(**Burt et al., 1976**)
Derived Parameters		
Parameter		Source
$[D1]$ in nM	≈ 1600	Eq.(17)
$[D2]$ in nM	≈ 80	Eq.(17)
$k_{on}^{D1,used}$ in $nm^{-1} min^{-1}$	0.0003125	see Text
$k_{off}^{D1,used}$ in min^{-1}	0.5	see Text

Table 1. Receptor parameters

The parameters that we used in the simulations are summarized in Tab. 1.

Dopamine signals

In our model we assumed a baseline [DA] of $[DA]^{tonic} = 20\ nM$ (**Dreyer et al., 2010**; **Dreyer, 2014**; **Venton et al., 2003**; **Suaud-Chagny et al., 1992**; **Borland et al., 2005**; **Justice Jr, 1993**; **Atcherley et al., 2015**). We modelled changes in [DA] to mimic DA signals observed in experimental studies. We use three types of single pulse DA signals: (long-)burst, burst-pause and ramp.

The burst signal mimics the result of a phasic burst in the activity of DA neurons in the SNc, e.g. in response to reward-predicting cues (**Pan et al., 2005**). The model burst signal consists of a rapid linear [DA] increase (with an amplitude $\Delta[DA]$ and rise time t_{rise}) and a subsequent return to baseline. The return to baseline is governed by Michaelis Menten kinetics with appropriate parameters for the dorsal striatum $V_{max} = 4.0\ \mu M s^{-1}$ and $K_m = 0.21\ \mu M$ (**Bergstrom and Garris, 2003**) and the nucleus accumbens $V_{max} = 1.5\ \mu M s^{-1}$

(*Dreyer and Hounsgaard, 2013*). In our model the removal of DA is assumed to happen without further DA influx into the system (baseline firing resumes when [DA] has returned to its baseline value). Unless stated otherwise, the long-burst signals are used with a $\Delta[DA] = 200\ nM$ and a rise time of $t_{rise} = 0.2\ s$ at $V_{max} = 1.5\ \mu M s^{-1}$, similar to biologically realistic transient signals (*Cheer et al., 2007*; *Robinson et al., 2001*; *Day et al., 2007*).

The burst-pause signal has two components, an initial short, small amplitude burst ($\Delta[DA] = 100\ nM$, $t_{rise} = 0.1\ s$), with the corresponding [DA] return to baseline (as for the long burst above). However, there is a second component in the DA signal, in which [DA] falls below baseline, simulating a pause in DA neuron firing. The length of this firing pause is characterized by the parameter t_{pause}. This burst-pause [DA] signal reflects the DA cell firing pattern consisting of a brief burst followed by a pause in activity (*Pan et al., 2008*; *Schultz, 2016*).

The ramp DA signal is characterized by the same parameters as the burst pattern, but with a longer t_{rise}, and a smaller $\Delta[DA]$.

Behavioural task simulation

To determine whether DA receptor occupancy can integrate reward signals over minutes, we simulated sequences consisting of 50 trials. Each sequence had a fixed reward probability. The trials contained either a long burst DA signal (mimicking a reward) or a burst-pause DA signal (mimicking no reward) at the beginning of the trial according to the reward probability of the sequence. The inter trial interval was $15 \pm 5s$ (**Fig. 2** and **Supp. Fig. 6**). We choose this highly simplistic scenario to reflect DA signals in a behavioural task in which the animal is rewarded for correct performance. However, here the specifics of the task are not relevant as our model addresses the integration of the DA receptor occupancy over time. Although we chose to use the burst-pause type signal as shown in **Fig. 1e** as a non-rewarding event, the difference to a non-signal are minimal after the end of the pause (**Supp. Figs. 2 and 5**). Each sequence started from a baseline receptor occupancy, assuming a break between sequences long enough for the receptors to return to baseline occupancy (around 5 minutes). For the simulations shown in **Supp. Fig. 5** all trials started exactly 15 s apart.

We simulated all reward probabilities from 0% to 100% in 10% steps. For each reward probability we ran 500 sequences, and calculated the mean receptor occupancy over time (single realisations shown in **Fig. 2a, b**). To investigate whether the receptor occupancy distinguished between different reward probabilities we applied a simple classifier to the receptor occupancy timeline.

The classifier was used to compare two different reward probabilities at a time. At each time point it was applied to a pair of receptor occupancies, e.g. one belonging to a 50% and one to a 30% reward probability sequence. The classifier assigned the current receptor occupancy to the higher or lower reward probability depending on which one was closer to the mean (over 500 sequences) receptor occupancy of that reward probability. As we knew the underlying reward probability of each sequence we were able to calculate the true and false positive rates and accuracy for each time point in our set of 500 sequences for both the D1R and D2R (**Supp. Fig. 6**). The accuracy was calculated based on all time points between 200 and 800s within a sequence to avoid the effect of the initial "swing-in" and post-sequence DA levels returning to baseline.

Acknowledgments

We thank Joshua Berke, Paul Overton, Alejandro Jimenez, Mohammadreza Mohagheghi Nejad and Amin Mirzaei for helpful discussions. This work was supported by the University of Sheffield and its high performance computing resources, the BrainLinks-BrainTools Cluster of Excellence funded by the German Research Foundation (DFG, grant number EXC 1086), and the state of Baden-Wuerttemberg through bwHPC.

References

Atcherley CW, Wood KM, Parent KL, Hashemi P, Heien ML. The coaction of tonic and phasic dopamine dynamics. Chemical Communications. 2015; 51(12):2235–2238.

Banay-Schwartz M, Kenessey A, DeGuzman T, Lajtha A, Palkovits M. Protein content of various regions of rat brain and adult and aging human brain. Age. 1992; 15(2):51–54.

Bergstrom BP, Garris PA. 'Passive stabilization'of striatal extracellular dopamine across the lesion spectrum encompassing the presymptomatic phase of Parkinson's disease: a voltammetric study in the 6-OHDA-lesioned rat. Journal of neurochemistry. 2003; 87(5):1224–1236.

Berke JD. What does dopamine mean? Nature neuroscience. 2018; p. 1.

Borland LM, Shi G, Yang H, Michael AC. Voltammetric study of extracellular dopamine near microdialysis probes acutely implanted in the striatum of the anesthetized rat. Journal of neuroscience methods. 2005; 146(2):149–158.

Burt DR, Creese I, Snyder SH. Properties of [3H] haloperidol and [3H] dopamine binding associated with dopamine receptors in calf brain membranes. Molecular pharmacology. 1976; 12(5):800–812.

Centonze D, Picconi B, Gubellini P, Bernardi G, Calabresi P. Dopaminergic control of synaptic plasticity in the dorsal striatum. European journal of neuroscience. 2001; 13(6):1071–1077.

Cheer JF, Aragona BJ, Heien ML, Seipel AT, Carelli RM, Wightman RM. Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. Neuron. 2007; 54(2):237–244.

Day JJ, Roitman MF, Wightman RM, Carelli RM. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nature neuroscience. 2007; 10(8):1020.

Day M, Wokosin D, Plotkin JL, Tian X, Surmeier DJ. Differential excitability and modulation of striatal medium spiny neuron dendrites. Journal of Neuroscience. 2008; 28(45):11603–11614.

DiResta G, Lee J, Lau N, Ali F, Galicich J, Arbit E. Measurement of brain tissue density using pycnometry. In: *Brain Edema VIII* Springer; 1990.p. 34–36.

Dreyer JK. Three mechanisms by which striatal denervation causes breakdown of dopamine signaling. Journal of Neuroscience. 2014; 34(37):12444–12456.

Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD. Influence of phasic and tonic dopamine release on receptor activation. Journal of Neuroscience. 2010; 30(42):14273–14283.

Dreyer JK, Hounsgaard J. Mathematical model of dopamine autoreceptors and uptake inhibitors and their influence on tonic and phasic dopamine signaling. Journal of neurophysiology. 2013; 109(1):171–182.

Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nature neuroscience. 2005; 8(11):1481.

384 **Flores-Barrera E**, Vizcarra-Chacón BJ, Bargas J, Tapia D, Galarraga E. Dopaminergic modulation of
385 corticostriatal responses in medium spiny projection neurons from direct and indirect pathways.
386 Frontiers in systems neuroscience. 2011; 5:15.

387 **Floresco SB**, West AR, Ash B, Moore H, Grace AA. Afferent modulation of dopamine neuron firing
388 differentially regulates tonic and phasic dopamine transmission. Nature neuroscience. 2003;
389 6(9):968.

390 **Frank MJ**, O'Reilly RC. A mechanistic account of striatal dopamine function in human cognition:
391 psychopharmacological studies with cabergoline and haloperidol. Behavioral neuroscience. 2006;
392 120(3):497.

393 **Grace AA**. The tonic/phasic model of dopamine system regulation: its relevance for understanding
394 how stimulant abuse can alter basal ganglia function. Drug & Alcohol Dependence. 1995;
395 37(2):111–129.

396 **Grace AA**, Floresco SB, Goto Y, Lodge DJ. Regulation of firing of dopaminergic neurons and control
397 of goal-directed behaviors. Trends in neurosciences. 2007; 30(5):220–227.

398 **Hamid AA**, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT,
399 Aragona BJ, Berke JD. Mesolimbic dopamine signals the value of work. Nature neuroscience.
400 2016; 19(1):117.

401 **Howe MW**, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM. Prolonged dopamine signalling in
402 striatum signals proximity and value of distant rewards. Nature. 2013; 500(7464):575–579.

403 **Justice Jr J**. Quantitative microdialysis of neurotransmitters. Journal of neuroscience methods.
404 1993; 48(3):263–276.

405 **Lamb T**, Pugh Jr E. G-protein cascades: gain and kinetics. Trends in neurosciences. 1992; 15(8):291–
406 298.

407 **Maeno H**. Dopamine receptors in canine caudate nucleus. Molecular and cellular biochemistry.
408 1982; 43(2):65–80.

409 **Marcott PF**, Mamaligas AA, Ford CP. Phasic dopamine release drives rapid activation of striatal
410 D2-receptors. Neuron. 2014; 84(1):164–176.

411 **Morris G**, Arkadir D, Nevet A, Vaadia E, Bergman H. Coincident but distinct messages of midbrain
412 dopamine and striatal tonically active neurons. Neuron. 2004; 43(1):133–143.

413 **Neve KA**, Neve RL. Molecular biology of dopamine receptors. In: *The dopamine receptors* Springer;
414 1997.p. 27–76.

415 **Nishikori K**, Noshiro O, Sano K, Maeno H. Characterization, solubilization, and separation of two
416 distinct dopamine receptors in canine caudate nucleus. Journal of Biological Chemistry. 1980;
417 255(22):10909–10915.

418 **Niv Y**, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response
419 vigor. Psychopharmacology. 2007; 191(3):507–520.

420 **Pan WX**, Schmidt R, Wickens JR, Hyland BI. Dopamine cells respond to predicted events during
421 classical conditioning: evidence for eligibility traces in the reward-learning network. Journal of
422 Neuroscience. 2005; 25(26):6235–6242.

423 **Pan WX**, Schmidt R, Wickens JR, Hyland BI. Tripartite mechanism of extinction suggested by
424 dopamine neuron activity and temporal difference model. Journal of Neuroscience. 2008;
425 28(39):9619–9631.

Patriarchi T, Cho JR, Merten K, Howe MW, Marley A, Xiong WH, Folk RW, Broussard GJ, Liang R, Jang MJ, et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. Science. 2018; p. eaat4422.

Prou D, Gu WJ, Le Crom S, Vincent JD, Salamero J, Vernier P. Intracellular retention of the two isoforms of the D 2 dopamine receptor promotes endoplasmic reticulum disruption. Journal of Cell Science. 2001; 114(19):3517–3527.

Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, Agid Y, DeLong MR, Obeso JA. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. Nature Reviews Neuroscience. 2010; 11(11):760–772.

Reyes BA, Pendergast JS, Yamazaki S. Mammalian peripheral circadian oscillators are temperature compensated. Journal of biological rhythms. 2008; 23(1):95–98.

Reynolds JN, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. Nature. 2001; 413(6851):67.

Richfield EK, Penney JB, Young AB. Anatomical and affinity state comparisons between dopamine D 1 and D 2 receptors in the rat central nervous system. Neuroscience. 1989; 30(3):767–777.

Richfield EK, Young AB, Penney JB. Comparative distribution of dopamine D-1 and D-2 receptors in the basal ganglia of turtles, pigeons, rats, cats, and monkeys. Journal of Comparative Neurology. 1987; 262(3):446–463.

Robinson DL, Phillips PE, Budygin EA, Trafton BJ, Garris PA, Wightman RM. Sub-second changes in accumbal dopamine during sexual behavior in male rats. Neuroreport. 2001; 12(11):2549–2552.

Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. Journal of Neuroscience. 2009; 29(42):13365–13376.

Roitman MF, Stuber GD, Phillips PE, Wightman RM, Carelli RM. Dopamine operates as a subsecond modulator of food seeking. Journal of Neuroscience. 2004; 24(6):1265–1271.

Roitman MF, Wheeler RA, Wightman RM, Carelli RM. Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. Nature neuroscience. 2008; 11(12):1376.

Sano K, Noshiro O, Katsuda K, Nishikori K, Maeno H. Dopamine receptors and dopamine-sensitive adenylate cyclase in canine caudate nucleus: Characterization and solubilization. Biochemical pharmacology. 1979; 28(24):3617–3627.

Schultz W. Predictive reward signal of dopamine neurons. Journal of neurophysiology. 1998; 80(1):1–27.

Schultz W. Multiple dopamine functions at different time courses. Annu Rev Neurosci. 2007; 30:259–288.

Schultz W. Dopamine reward prediction-error signalling: a two-component response. Nature Reviews Neuroscience. 2016; 17(3):183.

Suaud-Chagny M, Chergui K, Chouvet G, Gonon F. Relationship between dopamine release in the rat nucleus accumbens and the discharge activity of dopaminergic neurons during local in vivo application of amino acids in the ventral tegmental area. Neuroscience. 1992; 49(1):63–72.

Sun F, Zeng J, Jing M, Zhou J, Feng J, Owen SF, Luo Y, Li F, Wang H, Yamaguchi T, et al. A genetically encoded fluorescent sensor enables rapid and specific detection of dopamine in flies, fish, and mice. Cell. 2018; 174(2):481–496.

Surmeier DJ, Ding J, Day M, Wang Z, Shen W. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. Trends in neurosciences. 2007; 30(5):228–235.

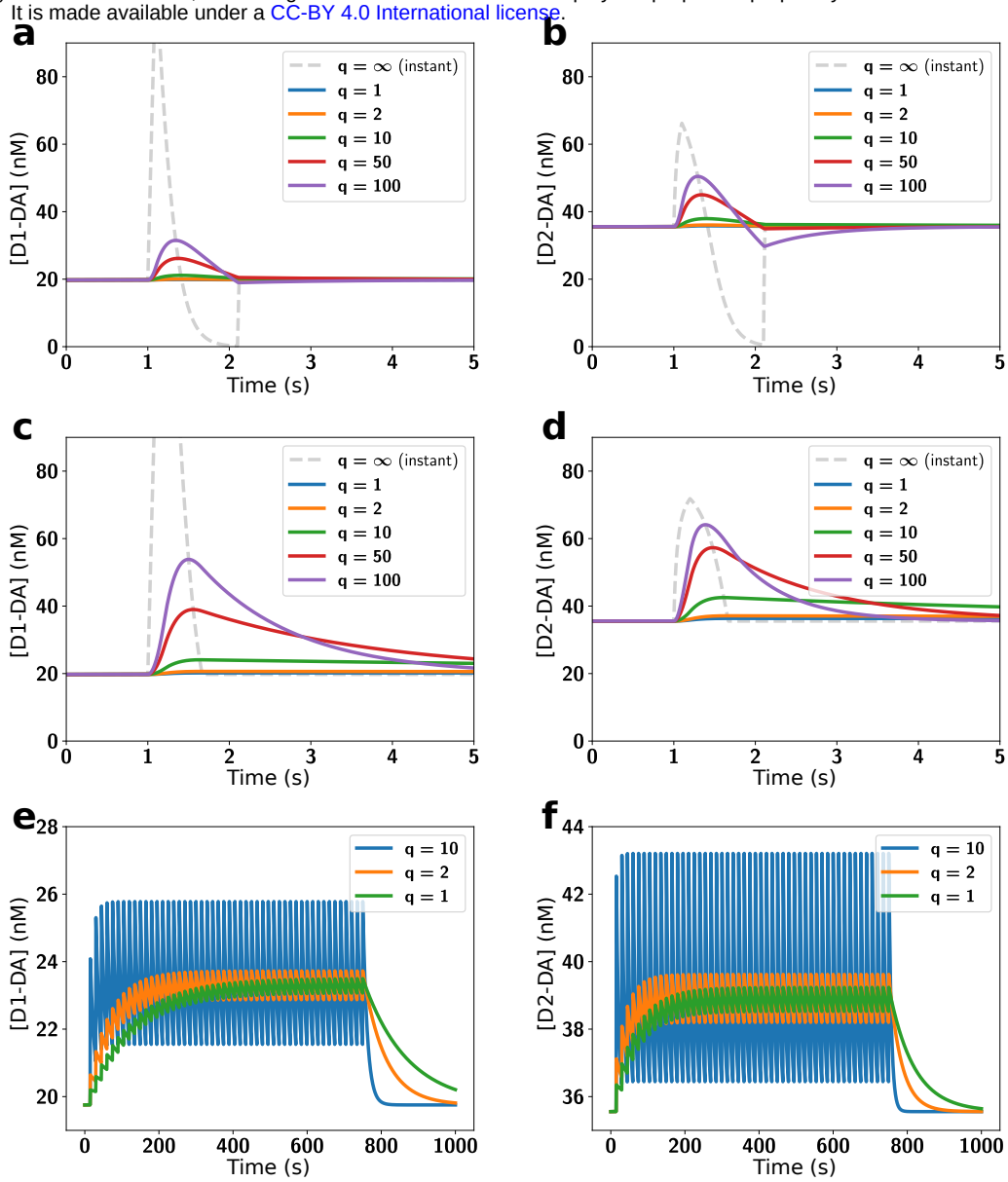
Syed EC, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME. Action initiation shapes mesolimbic dopamine encoding of future rewards. Nature neuroscience. 2016; 19(1):34.

Syková E, Nicholson C. Diffusion in brain extracellular space. Physiological reviews. 2008; 88(4):1277–1340.

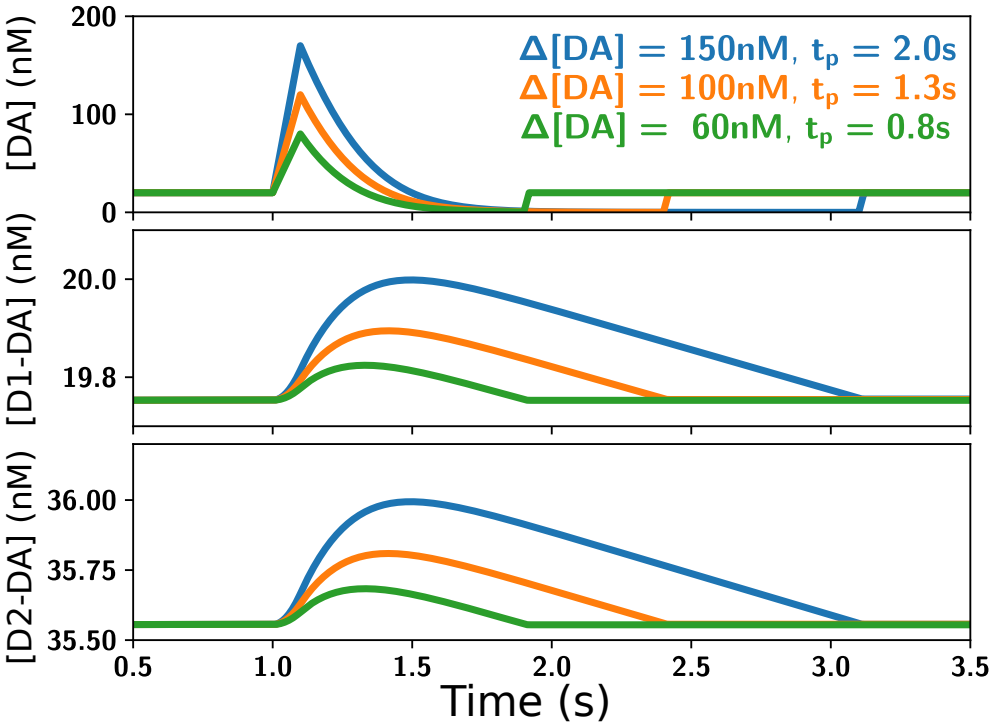
Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. Science. 2005; 307(5715):1642–1645.

Venton BJ, Zhang H, Garris PA, Phillips PE, Sulzer D, Wightman RM. Real-time decoding of dopamine concentration changes in the caudate–putamen during tonic and phasic firing. Journal of neurochemistry. 2003; 87(5):1284–1295.

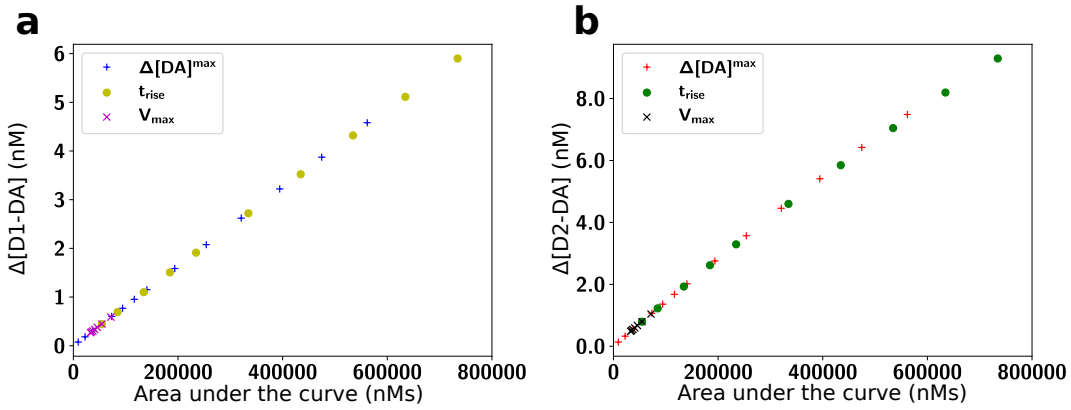
Yapo C, Nair AG, Clement L, Castro LR, Hellgren Kotaleski J, Vincent P. Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons. The Journal of physiology. 2017; 595(24):7451–7475.



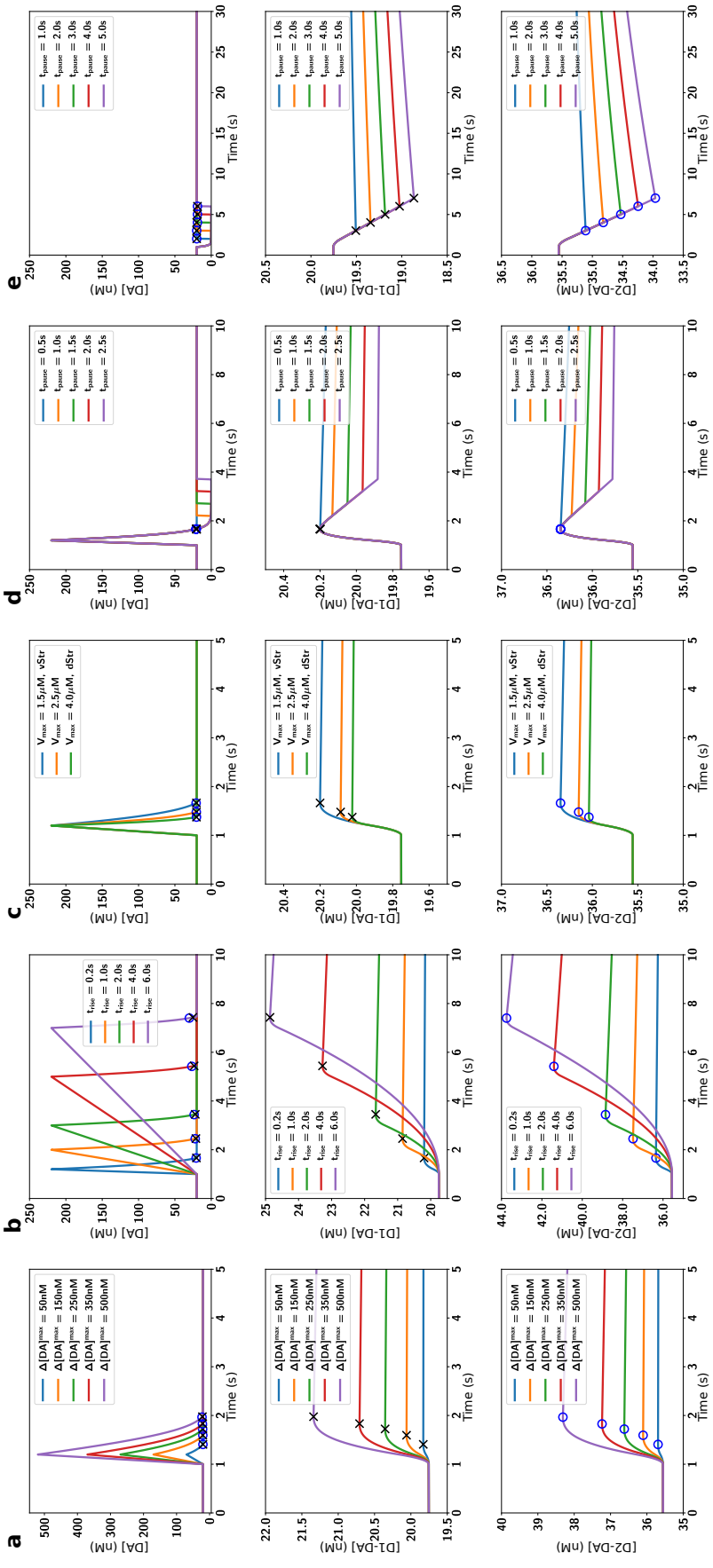
Supplemental Fig. 1. Similarities between D1 and D2 responses persist even if kinetics are much faster than our estimate. Fast kinetics were implemented by multiplying k_{on} and k_{off} by q as indicated, keeping K_D constant. Absolute D1R occupancy ([D1-DA]; left column) and D2R occupancy ([D2-DA]; right column) were examined for burst-pause DA signals **(a, b)**, burst-only DA signals **(c, d)**, and the behavioural sequence **(e, f)** (i.e. same simulation scenarios as in Fig. 1e and Supp. Fig. 5). D1Rs and D2Rs reacted very similarly to each other in all [DA] signal scenarios even if their kinetics were up to 100x faster because the difference between the aggregate D1 and D2 binding rates (Eq. 5) only differs by a factor of 1.5. Furthermore, the D2Rs do not show visible saturation effects even for $q = 100$. Faster kinetics mostly affected the amplitude of the receptor response and the time it takes to return to baseline receptor occupancy. However, only for $q = 100$ the pauses dropped slightly below baseline receptor occupancy (a, b). On a longer time scale with repetitive DA bursts (e, f) D1Rs and D2Rs integrated the DA bursts over time for $q = 1$ and $q = 2$. This is because the half-time of the receptors were 80 s (for $q = 1$), while the DA burst signal was repeated every 15 s. Thereby, [D1-DA] and [D2-DA] were dominated by the repetition of the signal rather than by the impact of individual DA burst signals. In contrast, for $q = 10$ the change in receptor occupancy was dominated by the single pulses, since the half-life time was 8s, whereby the receptors mostly unbind in between DA pulses.



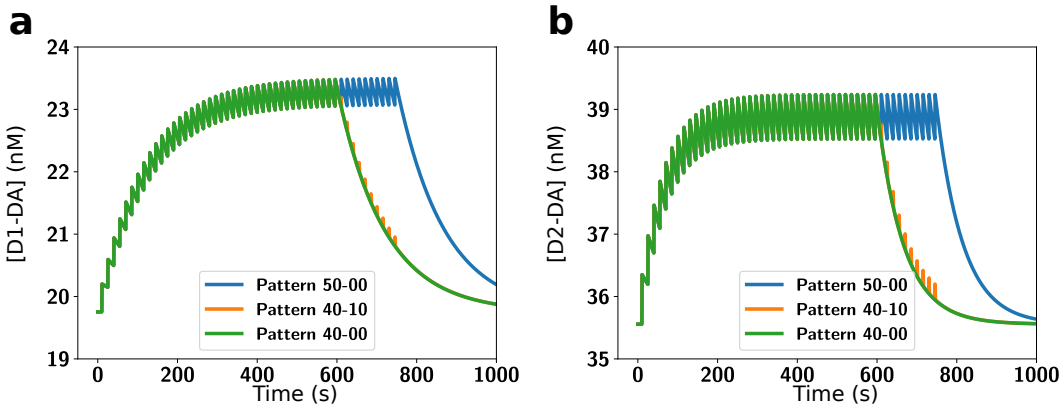
Supplemental Fig. 2. Burst-pause DA signals did not lead to increased DA receptor occupancy after the signal, if the duration of the pause was matched to the amplitude of the burst. Panels indicate input DA signal (top) and resulting DA receptor-ligand binding (middle and bottom). The return to baseline after the pulse happened for both D1R and D2R since their aggregate kinetics are similar.



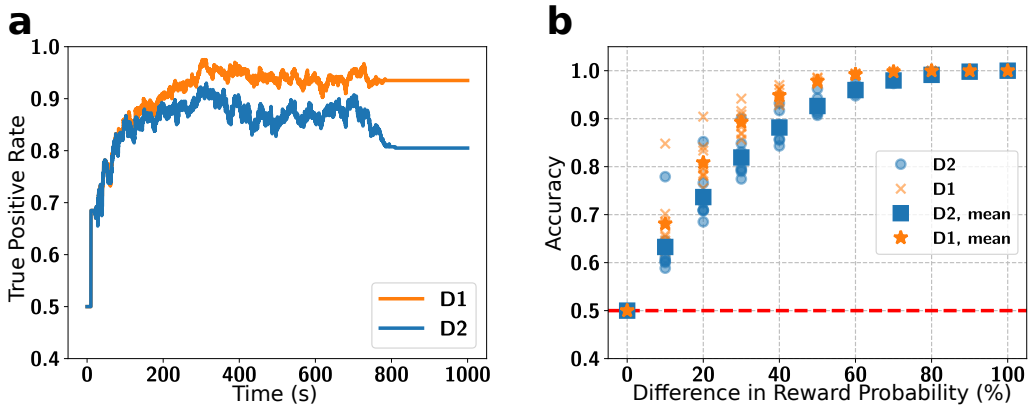
Supplemental Fig. 3. In the slow kinetics model the peak change in absolute receptor occupancy of D1Rs **(a)** and D2Rs **(b)** increased linearly with the area under the curve (AUC) of the DA pulses and parameter variations as in **Supp. Fig. 4a-c** (but with more parameter values). Here $\Delta[DA]^{max}$ marks burst-only DA pulses with varying peak amplitudes (50, 100, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000 nM), t_{rise} indicates ramping DA signals with with varying rise times (0.2, 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0 s) and V_{max} indicates burst-only DA pulses with varying V_{max} (1.0, 1.5, 2.0, 2.5, 3.0, 2.5, 4.0 in $\mu M\, s^{-1}$). This indicates that D1Rs and D2Rs act as slow integrators of the DA signal and explains why ramps are an effective signal to occupy DA receptors.



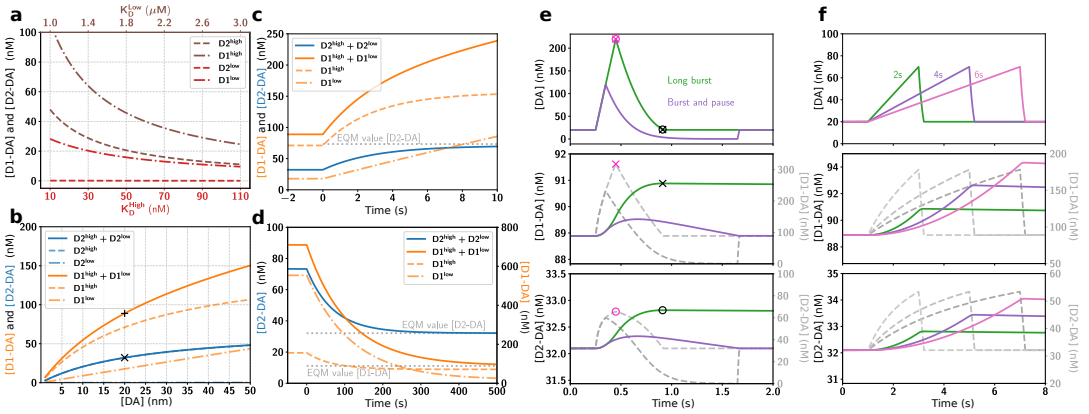
Supplemental Fig. 4. Parameter exploration for different DA signals (top row) with the resulting D1 (middle row) and D2 (bottom row) absolute receptor occupancy. The different input DA signals were burst-only with varying amplitude **(a)**, ramp with varying rise time **(b)**, burst-only with varying reuptake **(c)**, burst-pause with varying pause duration **(d)**, and pause-only with varying pause duration **(e)**. Blue circles and black crosses mark the time points of maximum receptor occupancy for D1 and D2, respectively (a-d), and of minimal receptor activation in (e). Note that for both D1 and D2 the time of maximum receptor occupancy was near the end of the DA signal and that D1Rs and D2Rs behaved similarly independent of the specific parameters of the DA pulse.



Supplemental Fig. 5. Absolute receptor occupancy for D1Rs **(a)** and D2Rs **(b)** following a sequence of 50 events. The sequences consists of 50 rewarding long burst events (blue), 40 long burst and 10 burst-pause events (orange) and 40 long burst events followed by 10 none events (green). A sequence of events lead to an accumulation of occupied receptors, if the pause between long burst events were shorter than $\approx 2 \cdot t_{1/2}$. There was a plateau for the absolute amount of occupied receptor at the level at which the amount of receptors unbinding until the next burst is the same as the amount of receptors getting occupied during a long burst event. Burst-pause events did not lead to an accumulation of occupied receptors over time and were, except during the short bursts, identical to the none events (note the overlapping green and orange curves), in line with a “false alarm" signal over a range of occupancy levels.



Supplemental Fig. 6. (a) True positive rates for the classification in a sample session (70% vs 30% reward probability) based on the receptor occupancy (see Methods) of D1 (orange) and D2 (blue) receptors. After a short 'swing-in" the receptors could distinguish between a 70% and a 30% reward rate. **(b)** Accuracy of the classifier for a range of reward probability differences for the D1 (orange) and D2 (blue) receptors (see Methods) for individual sessions and corresponding session averages.



Supplemental Fig. 7. Impact of slow kinetics on D1R and D2R binding with 10% of D1R in a high affinity state ($D1^{high}$) and 10% of D2 receptors in a low affinity state ($D2^{low}$) (***Richfield et al., 1989***). The $D1^{high}$ state was modelled by increasing the on-rate of the D1R but keeping its off-rate constant, creating a receptor identical to the $D2^{high}$ receptor. We choose this model since the high affinity state kinetics of the D1R are currently unknown, and a faster on-rate could potentially have the strongest effect on our conclusions. Correspondingly, we modelled the $D2^{low}$ receptor as a D2R with slower on-rate, which was largely equivalent to simply reducing $[D2^{tot}]$ since the $D2^{low}$ receptors were predominantly unoccupied during baseline DA and bound only sluggishly to DA during phasic signals. The main effect of incorporating the different receptor affinity states was a change in the respective equilibrium values of absolute concentration of receptors bound to DA. **(a)** The receptor occupancy at baseline $[DA] = 20nM$ was dominated by the high affinity states for both receptors, even though only 10% of the D1R were in the high state. As in our default model also D1 receptors were occupied at baseline, enabling them to detect tonic DA signals. **(b)** The amount of bound D1R and D2R stayed within the same order of magnitude over a range of baseline [DA]. '×' and +' indicate the model default parameters. **(c)** As in the default model, for a large step up from $[DA] = 20nM$ to $[DA] = 1\mu M$, and **(d)** a step down from $[DA] = 1\mu M$ to $[DA] = 20nM$, D1 and D2 receptor occupancy approached their new equilibrium (EQM, grey dotted lines) only slowly (i.e. over seconds to minutes). As the [D1-DA] changes were dominated by the $D1^{high}$ component, they were very similar to the D2R responses. **(e, f)** The effect of different phasic DA signals (top panels) was still very different in the slow kinetics model accounting for affinity states (coloured traces in middle and bottom panels; left scales) compared to the instant kinetics model (dashed grey traces, right scales). As in the default model, the timing of the maximum receptor occupancy ('×' and 'o' for D1 and D2, respectively) coincides for instant kinetics (purple symbols) with the [DA] peak (combined x and o in top panel), while for slow kinetics (black symbols) it coincides with the offset of the [DA] signal instead (combined x and o in top panel). The main difference to the default model is the higher occupancy of the D1R, caused by the $D1^{high}$ component. There is not a two-component unbinding since the $D1^{high}$ and $D1^{low}$ have similar off-rates, but differing on-rates. Overall, also for receptors with two affinity states, DA ramps are very effective in occupying the receptors.