

Dopamine neurons do not constitute an obligatory stage in the final common path for the evaluation and pursuit of brain stimulation reward

Ivan Trujillo-Pisanty^{1,□a}, Kent Conover¹, Pavel Solis¹, Daniel Palacios¹, Peter Shizgal^{1,*},

¹ Centre for Studies in Behavioural Neurobiology, Concordia University, Montreal, Québec, H4B 1R6, Canada

□a Current Address: Center for the Neurobiology of Addiction, Pain, and Emotion, Department of Anesthesiology and Pain Medicine, Department of Pharmacology, University of Washington, Seattle, WA, 98195, USA

* peter.shizgal@concordia.ca

Supporting information

Power-frequency trade-off

Fig S1. Response-rate versus pulse-frequency graph for rat Bechr14. The number of responses emitted per 2-min trial by an exemplar rat (Bechr29) is plotted as a function of pulse frequency and optical power.

Fig S2. Response-rate versus pulse-frequency graph for rat Bechr19.

Fig S3. Response-rate versus pulse-frequency graph for rat Bechr21.

Fig S4. Response-rate versus pulse-frequency graph for rat Bechr26.

Fig S5. Response-rate versus pulse-frequency graph for rat Bechr27.

Fig S6. Response-rate versus pulse-frequency graph for rat Bechr28.

The data for rat Bechr29 are shown in Fig 4 in the main text.

Time allocation as a function of reward strength and cost

Fig S7. Time allocation as a function of reward strength and cost for rat Bechr14. **A:** Time allocation as a function of pulse frequency (reward strength) in the vehicle (upright triangles) and drug (inverted triangles) conditions. **B:** Time allocation as a function of price (opportunity cost) in the vehicle (squares) and drug (diamonds) conditions. In the radial-sweep condition, the pulse frequency was decreased and the price decreased concurrently, in stepwise fashion, over consecutive trials. Time allocation is plotted as a function of pulse frequency in panel **C:** and as a function of price in panel **D:**. Data from the vehicle condition are represented by circles, whereas data from the drug condition are represented by Stars of David. The error bars represent 95% confidence intervals.

Fig S8. Time allocation as a function of reward strength and cost for rat Bechr19.

Fig S9. Time allocation as a function of reward strength and cost for rat Bechr21.

Fig S10. Time allocation as a function of reward strength and cost for rat Bechr26.

Fig S11. Time allocation as a function of reward strength and cost for rat Bechr27.

Fig S12. Time allocation as a function of reward strength and cost for rat Bechr28.

The data for rat Bechr29 are shown in Fig 5 in the main text.

Derivation of the reward-mountain model

Acronyms and symbols employed in this supporting-information file are defined in Tab S1.

Table S1. Definition of acronyms and symbols

Acronym or Symbol	Definition
a	price-sensitivity exponent
BSR	brain stimulation reward
c	chronaxie of the strength-duration function for pulses (pulse duration at which the threshold current is twice rheobase)
C	chronaxie of the strength-duration function for trains (train duration at which $F_{firing_{hm}}$ is twice ρ_{Π})
C_r	Conditioned reward value
ChR2	channelrhodopsin-2
d	pulse duration
D_{burst}	duration of stimulation-induced burst of firing in the directly activated neurons
D_{train}	train duration (leading edge of first pulse to leading edge of final pulse)
DAPI	4',6-diamidino-2-phenylindole
eICSS	electrical intracranial self-stimulation
f_a	function that determines the average reward rate derived from performance of alternate activities
f_D	function that relates the duration of the stimulation-induced burst of firing to the train duration
f_F	frequency-following function that relates the induced firing frequency to the pulse frequency
f_N	function that translates the pulse duration and current into the number of electrically recruited, directly stimulated neurons
f_p	subjective-probability function
f_P	subjective-price function
f_R	reward-growth function for eICSS or oICSS
f_T	behavioral-allocation function
f_U	function that translates the reward rate and subjective rate of exertion into a payoff
f_{ϕ}	subjective-effort function
F_{bend}	parameter governing the abruptness of the roll-off in frequency following
F_{firing}	firing frequency induced by electrical or optical stimulation
$F_{firing_{hm}}$	firing frequency that generates a reward of half-maximal intensity
$F_{firing_{max}}$	maximum firing frequency induced by electrical or optical stimulation
F_{hm}	notation for $F_{pulse_{hm}}$ used in previous eICSS papers
F_{hm}^*	shorthand notation for $F_{pulse_{hm}}^*$
F_{pulse}	pulse frequency in an electrical or optical pulse train

Table S1. Definition of acronyms and symbols

Acronym or Symbol	Definition
$F_{pulse_{hm}}$	pulse frequency required to drive reward intensity to half its maximum value
$F_{pulse_{hm}}^*$	estimated pulse frequency required to drive reward intensity to half its maximum value if frequency-following fidelity were perfect
F_{ro}	pulse frequency in the center of the roll-off region of the frequency-following function
g	exponent governing the steepness of reward-intensity growth
ICSS	intracranial self-stimulation
I	eICSS current
K_{aa}	constant scaling the value of alternate activities
K_{da}	constant representing the scaling of dopamine release
K_{ec_L}	effort-cost scalar for leisure activities
K_{ec_W}	effort-cost scalar for work
K_F	unit-translation constant to convert pulses s^{-1} into firings s^{-1} neuron $^{-1}$
K_{IS}	current-distance constant
K_{NS}	neuron-distance constant
K_{rg}	reward-intensity scalar
N	number of directly activated neurons
MFB	medial forebrain bundle
oICSS	optical intracranial self-stimulation
p_{obj}	objective probability that a reward will be delivered upon satisfaction of the response requirement
p_{sub}	subjective probability that a reward will be delivered upon satisfaction of the response requirement
P_e	notation for P_{obj_e} used in previous eICSS papers
P_e^*	shorthand notation for $P_{obj_e}^*$
P_{obj}	objective opportunity cost (“price”) of a stimulation train
P_{obj_e}	objective price at which $\hat{T} = 0.5$ when $\hat{R} = \hat{R}_{max}$
$P_{obj_e}^*$	estimated objective price at which $\hat{T} = 0.5$ and $\hat{R} = \hat{R}_{max}$ if frequency-following fidelity were perfect
P_{sub}	subjective opportunity cost (“price”) of a stimulation train
$\vec{P}_{sub_{\hat{T}=0.5}}$	vector of subjective prices that hold time allocation midway between its minimum and maximum values
$P_{sub_{bend}}$	parameter controlling the abruptness of the transition from the “blade” to the “handle” of the subjective-price function
$P_{sub_{min}}$	minimal subjective price
\dot{R}_{aa}	average rate of reward from performance of alternate (leisure) activities
$\hat{\dot{R}}_{aa}$	\dot{R}_{aa} normalized to vary between 0 and 1
R_{bsr}	peak reward intensity achieved over the course of a stimulation train
\hat{R}_{bsr}	R_{bsr} normalized to vary between 0 and 1 as R_{bsr} rises from 0 to $R_{bsr_{max}}$
$\hat{R}_{bsr_{max}}$	maximum normalized reward intensity
$\vec{\hat{R}}_{bsr_{\hat{T}=0.5}}$	vector of normalized reward intensities that hold time allocation midway between its minimum and maximum values
\dot{R}_{bsr}	rate of brain-stimulation reward; reward intensity divided by the subjective price paid to procure the pulse train
T	time allocation
\hat{T}	time allocation normalized to rise between 0 and 1 as T rises from T_{min} to T_{max}

Table S1. Definition of acronyms and symbols

Acronym or Symbol	Definition
T_{max}	maximal time allocation
T_{mid}	value of T midway between T_{min} and T_{max} ; $T_{mid} = \hat{T}_{0.5}$
T_{min}	minimal time allocation
TH	tyrosine hydroxylase
U_L	payoff from alternate (“leisure”) activities (a.k.a “everything else”)
U_W	payoff from brain stimulation reward
\hat{U}_W	U_W normalized to vary between 0 and 1
YFP	(enhanced) yellow fluorescent protein
ρ_I	rheobase of the strength-duration function for pulses
ρ_{Π}	rheobase of the strength-duration function for trains
$\dot{\phi}_{obj_L}$	average work rate entailed in performing leisure activities
$\dot{\phi}_{sub_L}$	average rate of subjective exertion entailed in performing leisure activities
$\dot{\hat{\phi}}_{sub_L}$	$\dot{\phi}_{sub_L}$ normalized to vary between 0 and 1
$\dot{\phi}_{obj_W}$	work rate entailed in holding down the lever
$\dot{\phi}_{sub_W}$	subjective rate of exertion entailed in holding down the lever

The reward-mountain model provides a framework for integrating the frequency-sweep, pulse-sweep, and radial-sweep data (e.g., Figure 5) in a unified 3D space and for interpreting the drug-induced changes in the position of the resulting 3D structure (the reward mountain). The following derivation of the model first extends earlier depictions that were developed in the context of eICSS studies [1–3] and then adapts the model to accommodate oICSS of midbrain dopamine neurons.

The reward-mountain model predicts time allocation given experimenter-controlled variables that determine the strength and cost of the rewarding stimulation. In most experiments carried out to date, the strength variable is the pulse frequency within a fixed-duration stimulation train, and the cost variable is the work time (opportunity cost, price) required to procure a stimulation train. The functional machinery that generates time-allocation values from these inputs is summarized in Tab S2.

Table S2. The functions composing the reward-mountain model. The shell \rightarrow core functions (left column) map the values of variables that are manipulated or controlled into inputs to the core of the model. The core functions (middle column) map these inputs into the payoffs from work and leisure activities, whereas the core \rightarrow shell function map the payoffs into the observed dependent variable: time allocation. The accompanying Matlab[®] Live Script illustrates how the listed functions are implemented. Please see Tab S1 for definitions of the symbols.

shell \rightarrow core	core	core \rightarrow shell
$F_{firing} = f_F(F_{pulse})$		
$N = f_N(I, d)$		
$D_{burst} = f_D(D_{train})$	$F_{firing_{hm}} = f_H(D_{burst}, N)$	
	$R_{bsr} = f_R(F_{firing}, F_{firing_{hm}})$	
$p_{sub} = f_p(p_{obj})$		
$P_{sub} = f_P(P_{obj})$		
$\dot{\phi}_{sub_W} = f_\phi(\dot{\phi}_{obj_W})$	$U_W = f_U\left(\frac{R_{bsr}}{P_{sub}}, p_{sub}, \dot{\phi}_{sub_W}\right)$	
$_ = f_a(_, _, \dots _)$	$\dot{R}_{aa} = f_R(_)$	
$\dot{\phi}_{sub_L} = f_\phi(\dot{\phi}_{obj_L})$	$U_L = f_U(\dot{R}_{aa}, \dot{\phi}_{sub_L})$	$T = f_T(U_W, U_L)$

As the table implies, we distinguish between the “shell” and “core” of the model. The shell consists of the variables that are observed (time allocation), manipulated (pulse frequency, price), and controlled (stimulation parameters held constant, physical work required to hold down the lever, affordances of the test environment [4]). The shell is displayed within the space defined by the observed and manipulated variables.

The core (middle column of Tab S2) consists of the functions that compute the intensity of the reward produced by the stimulation train and combine this value with the opportunity and effort costs to generate what we call “payoffs.” Parallel functions in the core compute the value of the alternate activities that compete with pursuit of the stimulation for the rat’s behavior. A set of functions (left column) provides the input to the core by mapping the manipulated and controlled variables into the quantities from which payoffs are derived.

A single function (right column), based on the generalized matching law [5], translates the payoffs generated in the core into the time-allocation values that are manifested in the shell.

Shell variables are objective, whereas core variables are inferred subjective quantities. The core functions are the bridge between the objective inputs that are manipulated or controlled to the observed objective output, time allocation; their form and parameters explain why the manipulated variables cause time allocation to vary in the manner observed in the experiment.

Shell \rightarrow core functions

The arguments of the first three shell \rightarrow core functions are the four parameters that define a fixed-frequency pulse train: the pulse frequency (F_{pulse}), current (I), pulse duration (d), and train duration (D_{train}). These functions relay to the core functions the stimulation-induced frequency of firing (F_{firing}), the number of activated neurons (N), and the duration of the burst of increased firing (D_{burst}).

f_F : In studies of eICSS, the experimenter typically varies the pulse frequency during a fixed-duration stimulation train in order to control the intensity of the electrically-induced reward. The frequency-following function labeled **f_F** maps the manipulated shell variable, F_{pulse} into the corresponding core variable: the induced frequency of firing in the directly-stimulated substrate, F_{firing} . In the case of eICSS of the MFB, this function has been estimated by

psychophysical means and shown to be roughly scalar up to very high pulse frequencies [6], well beyond the typical range of the $F_{pulse_{hm}}$ values that locate the reward-mountain shell in the space defined by the two independent variables. Solomon et al. [6] showed that the following function provides a good fit to the frequency-following data for eICSS of the MFB:

$$F_{firing} = f_F(F_{pulse}) = K_F \times F_{bend} \times \left(\ln \left[1 + e^{\frac{F_{ro}}{F_{bend}}} \right] - \ln \left[1 + e^{\frac{F_{ro} - F_{pulse}}{F_{bend}}} \right] \right) \quad (S1)$$

where

f_F = the frequency-following function

F_{bend} = parameter determining the abruptness of the roll-off in the frequency response; units: *unitless*

F_{firing} = induced firing rate in the first-stage neurons;
units: *firings s⁻¹ neuron⁻¹*

F_{pulse} = the pulse frequency; units: *pulses s⁻¹*

F_{ro} = the pulse frequency in the center of the roll-off region; units: *pulses s⁻¹*

K_F = unit-translation constant;
units: *firings pulse⁻¹ neuron⁻¹*

A plot of the frequency-following function is shown in Fig S13. The form of the function is the same as the one described by Solomon et al. [6]. The choice of parameters is described below in section Parameters of the frequency-following function for oICSS.

Fig S13. Assumed frequency-following function for optical stimulation of midbrain dopamine neurons.

The induced firing frequency is plotted as a function of the optical pulse frequency. Note that frequency-following fidelity is increasingly poor as the pulse frequency increases. At low values, the firing frequency falls only slightly short of the pulse frequency, but by 40 pulses s⁻¹, the firing frequency is only ~80% of the pulse frequency. The maximum induced firing frequency is 51.6 spikes s⁻¹

f_N : The F_N function translates the pulse duration and current into the number of electrically excited neurons. These two variables determine conjointly the boundary of the region in which the stimulation excites reward-related neurons. Holding these variables constant, as is the case in most eICSS studies entailing measurement of the reward mountain and in prior studies employing the curve-shift method [7–9], circumvents the need to make assumptions about the spatial distribution of the directly-stimulated neurons subserving the rewarding effect and about their excitability to extracellular stimulation.

Hawkins (cited in [10]) proposed that the number of directly-stimulated (“first-stage”) neurons subserving the rewarding effect is roughly proportional to the current, when pulse duration is held constant. Empirical studies [11,12] show that the current required to excite a given first-stage neuron varies roughly as a rectangular, hyperbolic function of the pulse duration. Thus,

$$N = f_N(d, I) = \frac{K_{NS}}{K_{IS}} \times \frac{I}{\rho_I \times \left(1 + \frac{c}{d}\right)} \quad (\text{S2})$$

where

$$\begin{aligned} c &= \text{chronaxie; pulse duration (units: ms) for which the threshold current is twice } \rho_I \\ d &= \text{pulse duration (units: ms)} \\ f_N &= \text{the first-stage recruitment function} \\ I &= \text{current (units: } \mu A) \\ K_{NS} &= \text{neuron-distance constant; units: } \textit{neurons mm}^{-2} \\ K_{IS} &= \text{current-distance constant; units: } \mu A \textit{ mm}^{-2} \\ N &= \text{number of activated first-stage neurons; units: } \textit{neurons} \\ \rho_I &= \text{threshold current required to excite a first-stage neuron using a pulse of infinite duration; units: } \mu A \end{aligned} \quad (\text{S3})$$

f_D : The third stimulation parameter that has been held constant in most reward-mountain and curve-shift studies is the duration of the stimulation train. The f_D function translates the duration set by the experimenter into the duration of the stimulation-induced increase in the activity of the directly-stimulated substrate. If firing is time-locked to the stimulation pulses, then the duration of this increase, D_{burst} , will equal the train duration, D_{train} . Thus, we assume that

$$D_{burst} = f_D(D_{train}) = D_{train} \quad (\text{S4})$$

where

$$\begin{aligned} D_{burst} &= \text{duration of the stimulation-induced increase in firing above baseline in the directly stimulated neurons subserving the reward effect} \\ D_{train} &= \text{time interval between the leading edges of the first and last pulse in a stimulation train} \\ f_D &= \text{the duration-mapping function} \end{aligned}$$

f_p : Psychophysical methods have been used to describe the subjective-probability function for BSR, f_p (lower-case subscript) [13]. This function returns the subjective probability that a reward will be delivered upon payment of the price set by the experimenter. Over the range, 0.5 - 1.0, this function was determined to be roughly scalar. In the present study and in all other studies that have entailed measurement of the reward mountain, delivery of the reward upon satisfaction of the response requirement is certain ($p_{obj} = 1$).

$$\begin{aligned} p_{sub} &= f_p(p_{obj}) = p_{obj} \\ &\text{for } 0.5 \leq p_{obj} \leq 1 \end{aligned} \quad (\text{S5})$$

where

$$\begin{aligned} f_p &= \text{the subjective-probability function} \\ p_{obj} &= \text{objective probability that reward will be delivered upon satisfaction of the response requirement} \\ p_{sub} &= \text{subjective probability that reward will be delivered upon satisfaction of the response requirement} \end{aligned}$$

f_P : Past measurements of the reward-mountain in eICSS studies employed the work time required to trigger delivery

of a stimulation train (the price) as the cost variable, and that practice is continued here. The function labeled f_P (upper-case subscript) maps the required work time, P_{obj} , into the corresponding subjective variable, P_{sub} . The form and parameters of this function for eICSS of the MFB have been estimated by psychophysical means [14]. We showed that the following psychophysical function more accurately describes the opportunity cost of rewarding brain stimulation than the identity function or functions based either on hyperbolic or exponential discounting:

$$P_{sub} = f_P(P_{obj}) = P_{sub_{min}} + P_{sub_{bend}} \times \ln \left(1 + e^{\left[\frac{P_{obj} - P_{sub_{min}}}{P_{sub_{bend}}} \right]} \right) \quad (S6)$$

where

- f_P = the subjective-price function
- P_{obj} = the "price" of a stimulation train: the cumulative time the lever must be depressed to trigger reward delivery; units: s
- P_{sub} = the subjective price of a stimulation train; units: s
- $P_{sub_{min}}$ = the minimum subjective price; units: s
- $P_{sub_{bend}}$ = a constant that controls the abruptness of the transition. from "blade" to "handle;" unitless

For a different view of the subjective-price function, please see [15].

At higher prices, the output of the subjective-price function defined by Eq S6 converges on its input: the subjective price becomes indistinguishable from the objective one. However, as the objective price is reduced below ~ 3 s, the subjective price deviates from the objective price and eventually approaches an asymptotic value: $P_{sub_{min}}$. This asymptote has been interpreted to arise from the reduction and eventual disappearance of competition between lever depression and competing activities, such as grooming, resting, and exploring; performance of these competing activities is no longer perceived as beneficial once the available time for their execution becomes sufficiently short.

A plot of the subjective-price function is shown in Fig S14. The parameters employed are the mean values determined by Solomon et al. [14].

Fig S14. Subjective opportunity-cost ("price") function. The function maps the objective opportunity cost (cumulative lever-depression time required to trigger reward delivery), P_{obj} , into its subjective equivalent P_{sub} . The form and parameters of this function are based on measurements by Solomon et al. [6].

The arguments of the first five shell \rightarrow core functions are all variables controlled directly by the experimenter: $\{F_{pulse}, d, I, D, p, P\}$. These six variables are transformed by the first five shell \rightarrow core functions into inputs to the core. The arguments of the remaining two shell \rightarrow core functions listed in Tab S2 are variables arising from features of the test environment that the experimenter attempts to hold constant: the rate of physical work entailed in holding down the lever or in performing activities that compete with pursuit of the rewarding stimulation.

f_ϕ : The effort cost of the reward is the subjective rate of exertion entailed in holding down the lever. We know of no psychophysical studies that reveal the form of this function. That said, we can assume that it accelerates steeply as the effort cost approaches the physical capabilities of the rat. The objective effort cost has been held constant in this and in prior studies carried out in the reward-mountain paradigm. We assume that the subjective effort cost did not co-vary with the manipulated variables. To help ensure this in the current study, the lever was withdrawn for 2 s following the triggering of a stimulation train so as to provide time for any interfering motoric consequences of the stimulation to dissipate.

We treat the subjective rate of exertion required to depress the lever as a constant defined by:

$$\dot{\phi}_{subW} = f_{\phi} \left(\dot{\phi}_{objW}, K_{\phi} \right) \quad (S7)$$

where

f_{ϕ} = subjective-effort function (form unknown)

$\dot{\phi}_{objW}$ = rate of physical work required to hold down the lever; units: $J s^{-1}$

$\dot{\phi}_{subW}$ = subjective rate of exertion required to hold down the lever
in units we call “oomphs” s^{-1}

K_{ϕ} = unit-conversion constant; units: *oomphs* J^{-1}

The dots over $\dot{\phi}_{objW}$ and $\dot{\phi}_{subW}$ signify that we define these quantities as rates. (No dots are placed over F_{firing} and F_{pulse} because doing so would be superfluous and potentially misleading. Frequencies are inherently rates over time (the first time derivative of the pulse or spike number). By omitting the dots, we wish to avoid confusion between these rates and their changes over time (the second derivative of the pulse or spike number).)

The effect of the drug, if any, on the rate of subjective exertion is defined as:

$$\dot{\phi}_{subWdrug} = \dot{\phi}_{subWvehicle} \times K_{ecW} \quad (S8)$$

where

K_{ecW} = proportional drug-induced change in the subjective rate of exertion required to hold down the lever; unitless

In the vehicle condition, K_{ecW} assumes an implicit value of one.

Activities such as grooming and exploring also entail performance of physical work. Thus, f_{ϕ} is also applied to these activities:

$$\dot{\phi}_{subL} = f_{\phi} \left(\dot{\phi}_{objL}, K_{\phi} \right) \quad (S9)$$

where

$\dot{\phi}_{objL}$ = Average rate of physical work required to perform alternate (“leisure”) activities; units: $J s^{-1}$.

$\dot{\phi}_{subL}$ = average subjective rate of exertion entailed in performance of leisure activities in *oomphs* s^{-1} .

K_{ϕ} = unit-conversion constant; units: *oomphs* J^{-1}

We assume that $\dot{\phi}_{subL}$ does not covary systematically with the independent variables.

As in the case of the subjective rate of exertion entailed in work, we allow for drug-induced modulation of the subjective rate of exertion entailed in performance of leisure activities.

$$\dot{\phi}_{subLdrug} = \dot{\phi}_{subLvehicle} \times K_{ecL} \quad (S10)$$

where

K_{ecL} = proportional drug-induced change in the subjective rate of exertion required to hold down the lever; unitless

In the vehicle condition, K_{ecL} assumes an implicit value of one.

f_a : We do not know which aspects of the leisure activities that compete with pursuit of BSR give rise to reward signals in the brain nor how these signals are encoded. That is why both the arguments and output of the seventh

shell \rightarrow core function (f_a) have been left blank. However, we do know that valuation of these activities influences the allocation of time to pursuit of experimenter-controlled rewards: enrichment of the test environment shifts allocation towards alternate activities and away from the experimenter-controlled reward [16]. Thus a function such as f_a must exist. We include f_a in the list of shell \rightarrow core functions for completeness and in recognition of this requirement.

Core functions

Core functions determine the reward rates produced by work (lever depression) and leisure (alternate) activities. These are combined with the associated effort costs to yield a pair of payoffs, which are then passed to the core \rightarrow shell function for translation into time allocation.

f_R : The core receives a set of spike trains as a result of the combined action of the first three shell \rightarrow core functions $\{f_F, f_N, f_D\}$, one spike train from every activated first-stage neuron. According to the counter model [10, 11, 17], the effects of the spike trains delivered by the individual first-stage neurons are summed, and thus, the intensity of the rewarding effect is determined by the product of the number of activated first-stage neurons and the rate at which they are fired by the stimulation train. This is why a Π symbol is used in the flow diagrams to represent the drive produced by a pulse train of fixed duration on the scalar at the input of the reward-growth function (Fig 3).

The logistic form of the reward-growth function was described originally in operant-matching studies carried out by Gallistel's group [18–20]. Shizgal [21] proposed the following expression for this function:

$$R_{bsr} = f_R[f_F(F_{pulse})] = \hat{R}_{bsr} \times K_{rg} = \left(\frac{F_{firing}^g}{F_{firing}^g + F_{firing_{hm}}^g} \right) \times K_{rg} \quad (S11)$$

where

- f_R = the reward-growth function
- $F_{firing_{hm}}$ = firing rate required to drive reward intensity to half its maximum value; units: $firings\ neuron^{-1}\ s^{-1}$
- $F_{firing_{hm}} = f_F(F_{pulse_{hm}})$
- $F_{pulse_{hm}}$ = pulse frequency required to drive reward intensity to half its maximum value; units: $pulses\ s^{-1}$
- g = the exponent that determines the steepness of reward-intensity growth as a function of pulse frequency
- K_{rg} = reward-growth scalar; units: *hedons*
- R_{bsr} = reward intensity produced by F_{firing} ; units: *hedons*
- \hat{R}_{bsr} = normalized reward intensity produced by a pulse frequency of F_{pulse} , which, in turn, produces a firing frequency of F_{firing} in each first-stage neuron; unitless. $0 \leq \hat{R}_{bsr} \leq 1$

According to Eq S11,

$$\hat{R}_{bsr} = \frac{F_{firing}^g}{F_{firing}^g + F_{firing_{hm}}^g} \quad (S12)$$

When frequency-following fidelity is sufficiently high, the normalized reward intensity (\hat{R}_{bsr}) will approach one at high pulse frequencies. The lower the value of the location parameter ($F_{firing_{hm}}$) and the higher the value of the reward-growth exponent (g), the easier this will be to achieve.

To accommodate the predicted rescaling of the input to the reward-growth function for oICSS by dopamine-transporter blockade, a scalar, K_{da} , is added to Eq S11 in the drug condition of the experiment:

$$\begin{aligned}
R_{bsr\,drug} &= \left[\frac{\left(F_{firing\,drug} \times K_{da} \right)^g}{\left(F_{firing\,drug} \times K_{da} \right)^g + \left(F_{firing\,hm\,vehicle} \right)^g} \right] \times K_{rg} \\
&= \left[\frac{\left(F_{firing\,drug} \right)^g}{\left(F_{firing\,drug} \right)^g + \left(\frac{F_{firing\,hm\,vehicle}}{K_{da}} \right)^g} \right] \times K_{rg}
\end{aligned} \tag{S13}$$

where

K_{da} = scalar representing the boost in dopamine release due to transporter blockade

Thus,

$$F_{firing\,hm\,drug} = \frac{F_{firing\,hm\,vehicle}}{K_{da}} \tag{S14}$$

Dopamine-transporter blockade boosts dopamine release, thereby increasing the impact of each firing. The increased value of the scalar (K_{da}) captures this augmented impact, reducing the value of the location parameter of the reward-growth function in the drug condition. Fewer pulses per train are required to produce a reward of a given intensity when K_{da} increases. Consequently, the reward-growth function shifts leftwards along the pulse-frequency axis. In the simulations of the vehicle condition, K_{da} is assigned an implicit value of one.

Division by the subjective price transforms the reward intensity into a reward rate:

$$\dot{R}_{bsr} = \frac{R_{bsr}}{P_{sub}} \tag{S15}$$

where

\dot{R}_{bsr} = experienced rate of brain stimulation reward; units: $hedons\ s^{-1}$

f_H : The location parameter of the reward-growth function is the firing rate that drives reward intensity to half its maximal value. The value of this parameter depends on the number of stimulated first-stage neurons, N and the interval during which the stimulation train elevates their firing rate, D_{burst} . A prior study of temporal integration in the neural circuitry responsible for eICSS of the MFB [21] implies the following form for the function that determines $F_{firing\,hm}$:

$$F_{firing\,hm} = f_H(D_{burst}, N) = \frac{\rho_{\Pi} \times \left(1 + \frac{C}{D_{burst}} \right)}{N} \tag{S16}$$

where

C = chronaxie: train duration at which $F_{firing\,hm}$ is twice the value of ρ_{Π} ; units: s

ρ_{Π} = aggregate rate of firing required to produce a reward of half-maximal intensity when the train duration is infinite; units: $firings\ s^{-1}$

f_U : In keeping with the generalized matching law [22], the benefit from work and its costs are combined in scalar fashion to yield a net payoff. We can format the expression for the payoff as a ratio of two rates: the

probability-weighted rate of brain-stimulation reward and the subjective rate of exertion required to hold down the lever (see [23]):

$$U_W = f_U \left(\frac{R_{bsr}}{P_{sub}}, p_{sub}, \dot{\phi}_{sub_W} \right) = \frac{\frac{R_{bsr}}{P_{sub}} \times p_{sub}}{\dot{\phi}_{sub_W}} = \frac{\dot{R}_{bsr} \times p_{sub}}{\dot{\phi}_{sub_W}} \quad (S17)$$

where

f_U = utility function

U_W = payoff from a train of rewarding stimulation;
units: *hedons oomph*⁻¹

or as a benefit/cost ratio:

$$U_W = f_U \left(\frac{R_{bsr}}{P_{sub}}, p_{sub}, \dot{\phi}_{sub_W} \right) = \frac{R_{bsr} \times p_{sub}}{P_{sub} \times \dot{\phi}_{sub_W}} \quad (S18)$$

In the case of BSR, there is solid evidence that aggregate impulse flow in the first-stage neurons encodes the signal that will be translated into the intensity of the reward [17, 20]. Although the identity of the first-stage neurons subserving eICSS of the MFB (or any other brain site) remains unknown, directly-driven MFB neurons with properties that match the psychophysically derived portrait of the first-stage fibers have been observed by means of electrophysiological recording [24–26].

The argument of the reward-growth function for leisure activities is left blank, thus signifying our ignorance of how pursuit of these activities is encoded by the brain and translated into a reward rate. That said, the considerable evidence that the value of leisure activities competes effectively with experimenter-controlled rewards [5, 16, 27–29] implies that the payoffs from work and leisure are commensurable. Accordingly, we define a reward rate for the leisure activities that compete with pursuit of BSR:

$$\dot{R}_{aa} = f_R(_) = \dot{\hat{R}}_{aa} \times K_{aa} \quad (S19)$$

where

K_{aa} = alternate-activity scalar; units: *hedons s*⁻¹

\dot{R}_{aa} = average rate of reward from alternate (leisure) activities
that compete with pursuit of BSR; units: *hedons s*⁻¹

$\dot{\hat{R}}_{aa}$ = normalized rate of reward from alternate (leisure) activities;
unitless. $0 \leq \dot{\hat{R}}_{aa} \leq 1$

$(_) =$ the unknown variables that give rise to \dot{R}_{aa}

The payoff from these alternate activities is computed in a manner analogous to the computation of the payoff from BSR, as a ratio of reward and subjective-exertion rates:

$$U_L = f_U \left(\dot{R}_{aa}, \dot{\phi}_{sub_L} \right) = \frac{\dot{\hat{R}}_{aa} \times K_{aa}}{\dot{\phi}_{sub_L}} \quad (S20)$$

where

U_L = payoff from leisure activities;
units: *hedons oomph*⁻¹

The core \rightarrow shell function

f_T : The payoffs from pursuit of BSR (U_W) and engagement in leisure activities (U_L) are used by the sole core \rightarrow shell function to compute the allocation of time to pursuit of BSR. This behavioral-allocation function is derived from the single-operant matching law [5,30,31]:

$$T = T_{min} + \left[(T_{max} - T_{min}) \times \frac{(U_W)^a}{(U_W)^a + (U_L)^a} \right] \quad (\text{S21})$$

where

a = price-sensitivity exponent; unitless

T = time allocation; unitless

T_{max} = maximum time allocation; unitless

T_{min} = minimum time allocation; unitless

U_W = payoff from pursuit of rewarding brain stimulation
("work"); units: *hedons oomph*⁻¹

U_L = payoff from pursuit of alternate ("leisure") activities;
units: *hedons oomph*⁻¹

Even when the rat is working maximally, latency to depress the lever is typically greater than zero, and T_{max} is thus typically less than one. The rat tends to sample the lever at trial onset, even when the payoff from brain stimulation is low. Thus, T_{min} is typically greater than zero.

The single-operant matching law was formulated initially to account for the behavior of subjects working on variable-interval schedules of reinforcement. The cumulative handling-time schedule in force in the present study [32] is more akin to a fixed-ratio schedule in that the number of rewards earned is strictly proportional to time worked. On such schedules, time allocation shifts more abruptly than on variable-interval schedules as the value of the experimenter-controlled reward is varied. In the original version of the single-operant matching law [30,31], the payoff terms are not exponentiated. In contrast, the price-sensitivity exponent (a) in Eq S21 allows the reward-mountain model to account for time-allocation shifts of varying abruptness.

To simplify the remaining derivation of the reward-mountain model, we define a normalized measure of time allocation:

$$\hat{T} = \frac{(U_W)^a}{(U_W)^a + (U_L)^a} \quad (\text{S22})$$

$$\text{where } 0 \leq \hat{T} \leq 1$$

Substituting from Eq S22 in Eq S21, we obtain

$$T = T_{min} + \left[(T_{max} - T_{min}) \times \hat{T} \right] \quad (\text{S23})$$

$$\hat{T} = \frac{T - T_{min}}{T_{max} - T_{min}} \quad (\text{S24})$$

Time allocation as a function of reward strength and cost

We are now in a position to tie the dependent variable, time allocation (T), to the two independent variables, pulse frequency (F_{pulse}) and price (P_{obj}). Substitution for U_W and U_L from Eqs S17 and S20, respectively, in Eq S22 yields:

$$\hat{T} = \frac{\left(\frac{[\hat{R}_{bsr} \times K_{rg}] \times p_{sub}}{[\hat{\phi}_{subW} \times K_{ecW}] \times P_{sub}} \right)^a}{\left(\frac{[\hat{R}_{bsr} \times K_{rg}] \times p_{sub}}{[\hat{\phi}_{subW} \times K_{ecW}] \times P_{sub}} \right)^a + (U_L)^a} \quad (S25)$$

An initial step towards simplifying Eq S25 is to multiply each of the terms on the right by

$$\left(\frac{K_{ecW}}{K_{rg}} \right) \times \left(\frac{\hat{\phi}_{subW} \times P_{sub}}{p_{sub}} \right)$$

This yields:

$$\hat{T} = \frac{(\hat{R}_{bsr})^a}{(\hat{R}_{bsr})^a + \left[\left(\frac{K_{ecW} \times K_{aa}}{K_{rg}} \right) \times \left(\frac{\hat{\phi}_{subW} \times (U_L)^a}{p_{sub}} \right) \times P_{sub} \right]^a} \quad (S26)$$

To simplify Eq S25 further, we first define T_{mid} as the time-allocation value midway between maximal and minimal time allocation:

$$T_{mid} = T_{min} + \left(\frac{T_{max} - T_{min}}{2} \right) \quad (S27)$$

According to Eqs S22 and S23,

$$\text{when } T = T_{mid}$$

$$\hat{T} = 0.5$$

and

$$U_W = U_L \quad (S28)$$

We now hold time allocation at T_{mid} and drive reward intensity to its maximum value ($\hat{R}_{bsr} = \hat{R}_{bsr,max}$). Substituting for U_W from Eq S17 and reversing Eq S28, we obtain

$$U_L = \frac{[\hat{R}_{bsrmax} \times K_{rg}] \times p_{sub}}{[\dot{\hat{\phi}}_{sub_W} \times K_{ec_W}] \times P_{sub_e}(P_{obj_e})} \quad (S29)$$

where

P_{obj_e} = objective price at which $T = T_{mid}$ when $\hat{R}_{bsr} = \hat{R}_{bsrmax}$

P_{sub_e} = subjective price at which $T = T_{mid}$ when $\hat{R}_{bsr} = \hat{R}_{bsrmax}$

Rearranging the terms yields

$$P_{sub_e} = \frac{K_{rg}}{K_{aa} \times K_{ec_W}} \times \frac{p_{sub} \times \hat{R}_{bsrmax}}{\dot{\hat{\phi}}_{sub_W} \times U_L} \quad (S30)$$

For consistency with previous papers, we use the subscript, “e” in the symbol P_{sub_e} to refer to the fact that the payoff from brain stimulation equals the payoff from alternate activities (“everything else”) when the normalized reward intensity is maximal ($\hat{R}_{bsr} = \hat{R}_{bsrmax}$) and the subjective price (P_{sub}) equals P_{sub_e} . This equivalence between the two competing payoffs is what drives time allocation to the half-way point (T_{mid}) between its minimal (T_{min}) and maximal (T_{max}) values.

Rearranging Eq S30, we obtain:

$$\frac{\hat{R}_{bsrmax}}{P_{sub_e}} = \left[\left(\frac{K_{aa} \times K_{ec_W}}{K_{rg}} \right) \times \left(\frac{\dot{\hat{\phi}}_{sub_W} \times U_L}{p_{sub}} \right) \right] \quad (S31)$$

Substituting in Eq S26 from Eqs S11, S24, and S31, we obtain

$$\hat{T} = \frac{(\hat{R}_{bsr})^a}{(\hat{R}_{bsr})^a + \left[\hat{R}_{bsrmax} \times \left(\frac{P_{sub}}{P_{sub_e}} \right) \right]^a} \quad (S32)$$

By setting $\hat{R}_{bsr} = \hat{R}_{bsrmax}$ and $P_{sub} = P_{sub_e}$ in Eq S32, it can be seen readily that $\hat{T} = 0.5$, thus satisfying the definition of P_{sub_e} as the price at which time allocation to pursuit of a maximally intense reward is halfway between T_{min} and T_{max} .

P_{sub} and \hat{R}_{bsr} must trade off to hold \hat{T} at a given level. The lowest attainable subjective price is the value corresponding to an objective price of zero, which we will call P_{sub_0} . (Negative values of P_{obj} may be required to drive P_{sub} to P_{submin} .) Consider a vector of subjective prices, $\vec{P}_{sub_{\hat{T}}}$, that extends from P_{sub_0} to the highest tested value of P_{sub} and a corresponding vector of normalized reward intensities, $\vec{\hat{R}}_{bsr_{\hat{T}}}$, that hold normalized time allocation (\hat{T}) constant over \vec{P}_{sub} . It follows from Eq S32 that

$$\frac{\vec{R}_{bsr\hat{T}}}{\vec{P}_{sub\hat{T}}} = \left(\frac{\hat{T}}{1 - \hat{T}} \right)^{\frac{1}{a}} \times \left(\frac{\hat{R}_{bsrmax}}{P_{sube}} \right) \quad (S33)$$

$$0 \leq \vec{R}_{bsr\hat{T}} \leq \hat{R}_{bsrmax} \leq 1$$

$$0 < P_{sub0} \leq P_{sub\hat{T}}$$

where

P_{sub0} = the subjective price corresponding to an objective price of zero

The higher the subjective price, the higher the normalized reward intensity required to hold time allocation constant. 206

When $\hat{T} = 0.5$, Eq S33 reduces to: 207

$$\frac{\vec{R}_{bsr\hat{T}=0.5}}{\vec{P}_{sub\hat{T}=0.5}} = \frac{\hat{R}_{bsrmax}}{P_{sube}} \quad (S34)$$

$$0 \leq \hat{R}_{bsr\hat{T}=0.5} \leq \hat{R}_{bsrmax} \leq 1$$

$$0 < P_{sub0} \leq P_{sub\hat{T}=0.5} \leq P_{sube}$$

Below (see: Contour lines: the trade-off between pulse frequency and price to hold time allocation constant), we use Eqs S33 and S34 to obtain the equation for the contour lines that provide a two-dimensional description of the reward-mountain surface. 208
209
210

To complete the derivation of the reward-mountain model, we now substitute for \hat{T} in Eq S23 and expand Eq S32 so that time allocation is expressed in terms of the independent variables: price (P_{obj}) and pulse frequency (F_{pulse}), which appear as the arguments of the subjective-price and frequency-following functions, respectively: 211
212
213

$$T = T_{min} + (T_{max} - T_{min}) \times \frac{\left(\frac{f_F(F_{pulse})^g}{f_F(F_{pulse})^g + f_F(F_{pulse_{hm}})^g} \right)^a}{\left(\frac{f_F(F_{pulse})^g}{f_F(F_{pulse})^g + f_F(F_{pulse_{hm}})^g} \right)^a + \left(\hat{R}_{bsrmax} \times \left[\frac{f_P(P_{obj})}{f_P(P_{obj_e})} \right] \right)^a} \quad (S35)$$

The conditioned-reward variant of the reward-mountain model 214

The six-parameter version of the reward-mountain model incorporates Eqs S1, S6, and S35. The fitted parameters are a (the price-sensitivity exponent), $F_{pulse_{hm}}$ (the pulse frequency at which reward intensity is half maximal), g (the reward-growth exponent) P_{obj_e} (the price at which time allocation to pursuit of a maximal reward falls midway between its minimal and maximal values), T_{min} (minimum time allocation), and T_{max} (maximal time allocation). 215
216
217
218
The seven-parameter version includes an additional parameter, C_r , to reflect conditioned reward. This seventh parameter reflects a learned value, above and beyond the payoff from the stimulation train, associated with the lever and/or the act of holding it down. The paper that introduced this parameter [3] incorporated it into the reward-growth function as follows: 219
220
221
222

$$\hat{R}_{bsr} = \frac{F_{pulse}^g + Cr^g}{F_{pulse}^g + F_{pulse_{hm}}^g + Cr^g} \quad (S36)$$

where

C_r = the conditioned reward, expressed in terms of the equivalent pulse frequency

Note that the way the C_r parameter was incorporated into Eq S36 causes this parameter to interact with both the a (Eq S35) and g (Eq S36) parameters. That form of the model failed to yield consistently converging fits when applied to the current dataset. To address this problem, we altered the way that the C_r parameter is incorporated into the reward-growth function (Eq S11) so as to reduce its interaction with the other parameters:

$$\hat{R}_{bsr} = C_r + \left[\left(1 - \frac{C_r}{\hat{R}_{bsr_{max}}} \right) \times \left(\frac{f_F(F_{pulse})^g}{f_F(F_{pulse})^g + f_F(F_{pulse_{hm}})^g} \right) \right] \quad (S37)$$

The resulting reward-mountain surface is produced by substituting the expression for \hat{R}_{bsr} from Eq S37 in Eq S32, as follows:

$$T = T_{min} + (T_{max} - T_{min}) \times \frac{\left(C_r + \left[\left(1 - \frac{C_r}{\hat{R}_{bsr_{max}}} \right) \times \left(\frac{f_F(F_{pulse})^g}{f_F(F_{pulse})^g + f_F(F_{pulse_{hm}})^g} \right) \right] \right)^a}{\left(C_r + \left[\left(1 - \frac{C_r}{\hat{R}_{bsr_{max}}} \right) \times \left(\frac{f_F(F_{pulse})^g}{f_F(F_{pulse})^g + f_F(F_{pulse_{hm}})^g} \right) \right] \right)^a + \left(\hat{R}_{bsr_{max}} \times \left[\frac{f_P(P_{obj})}{f_P(P_{obj_e})} \right] \right)^a} \quad (S38)$$

As shown in Fig S15, Eq S38 yields a reward-mountain surface that is all but indistinguishable from the surface generated by the equation in the 2010 paper. Unlike the equation in the 2010 paper, Eq S38 produced well-behaved, converging fits.

Fig S15. Surface and contour plots of two seven-parameter reward-mountain models. The new version of the model produces a surface that is nearly identical to the one generated by the 2010 version of this model [3]. Fits of the 2010 model to the current datasets failed to converge, whereas fits of the new version of the model converged in all cases.

Adaptation of the reward-mountain model for oICSS

Fig S16. The reward-mountain model for oICSS. A graphical summary of the reward-mountain model, as adapted for oICSS of midbrain dopamine neurons. The symbols are defined in Tab S1.

The reward-mountain surface shows the observed behavioral output (time allocation, T) as a function of the two independent variables, the price (P_{obj}) and pulse frequency (F_{pulse}). These three variables constitute the shell of the reward-mountain model, together with the controlled variables: the work rate required to hold down the lever ($\dot{\phi}_{obj_W}$), the leisure activities afforded by the test environment, and the average work rate required to perform these activities ($\dot{\phi}_{obj_L}$). The shell \rightarrow core functions, $\{f_F, f_P, f_P, f_\phi, f_a\}$, map the manipulated and controlled variables into the core quantities that determine the payoffs from work and leisure. These mapping functions are shown at the left of the figure. The core functions that compute the reward intensities and corresponding payoffs $\{f_R, f_U\}$ are shown in the center. (f_H , the function that determines the location parameter of the reward-growth function, is not shown, nor are f_D , the function that determines the duration of the stimulation-induced burst of firing and f_N , the

function that determines the number of recruited neurons (N).) The core \rightarrow shell function, (f_T), translates the payoffs from work and leisure into time allocation. It is shown on the right.

Minimal changes were made to adapt the model prior to fitting the reward-mountain surface to the data:

1. New values were used for the parameters of the frequency-following function (Eq S1) so as to accommodate the known properties of midbrain dopamine neurons and the optical-power versus pulse-frequency trade-off data reported here.
2. In eICSS, the current and pulse duration conjointly determine the number of directly stimulated neurons. In oICSS, optical power plays the role assumed by current in eICSS. However, there is insufficient information available to model f_N , the function that determines the number of neurons directly activated by a given optical power and pulse duration. Instead, we simply chose values of N and ρ_{Π} that produce simulated reward-mountain surfaces similar to the ones returned by the fits of the model to the data. (See the accompanying Matlab[®] Live Script.)
3. $\hat{R}_{bsr_{max}}$ was included explicitly in the reward-mountain model, thus expanding the model to accommodate imperfect frequency-following fidelity.
4. Explicit inclusion of $\hat{R}_{bsr_{max}}$ required correction of the location parameters and a more nuanced treatment of drug-induced shifts in the location of the mountain. (See: Displacement of the shell: distinguishing two sources.)
5. The new conditioned-reward variant described above (Eq S38) was used in lieu of the version introduced in our prior studies of the effect of cocaine and GBR-12909 on reward mountains obtained in the eICSS paradigm [3,33].

Parameters of the frequency-following function for oICSS

The form of the frequency-following function (f_F) for channelrhodopsin-2-mediated excitation of midbrain dopamine neurons has yet to be determined. Here, we used a function of the same form as the one we had determined previously for eICSS of the MFB [6], but we substituted new parameter values. Dopamine neurons cannot fire nearly as fast as the directly stimulated neurons subserving the rewarding effect of electrical MFB stimulation [6,34,35], and the kinetics of channelrhodopsin-2 are slow in comparison to those of the voltage-gated channels responsible for electrically induced neural firing [36]. Thus, frequency-following parameters determined for eICSS of the MFB cannot be used to account for the frequency response of the dopamine neurons subserving oICSS of the ventral midbrain.

Although two studies found that frequency-following fidelity in optically stimulated midbrain dopamine neurons fell to only 40-50% at optical pulse frequencies of 40-50 pulses s^{-1} [34,35], results of two other electrophysiological studies show very good firing fidelity at 50 pulses s^{-1} [37,38]. Moreover, results of a recent electrochemical study [39] show that optically induced dopamine release in the nucleus accumbens continued to rise as the pulse frequency was increased from 40-50 pulses s^{-1} . Poor frequency-following fidelity was found by Lohani and colleagues in midbrain dopamine neurons optically stimulated at 100 pulses s^{-1} [40], suggesting that the upper limit on the induced firing rate lies at a significantly lower pulse frequency.

Figs 4 and S5 show two cases (data from rats BeChR29 and 27) in which the behavioral effectiveness of the stimulation continues to rise at pulse frequencies up to, or beyond, 60 pulses s^{-1} . The curves for the remaining rats approach asymptote earlier, but this does not necessarily reflect failure the higher pulse frequencies to increase firing: saturation of reward-intensity growth [20] could be responsible instead.

In view of the power-frequency trade-off data reported here and the results reported in the studies cited above, we set the middle of the roll-off region of the frequency-following function (F_{ro}) to 50 pulses s^{-1} and the parameter governing the abruptness of the roll-off (F_{bend}) to 20. The resulting frequency-following function is shown in Fig S13. It continues to climb up to optical pulse frequencies of ~ 100 pulses s^{-1} to attain a maximum induced firing rate of 51.6 spikes s^{-1} .

Displacement of the shell: distinguishing two sources

The purpose of the experiment is to draw inferences about brain-reward circuitry from the effect of dopamine-transporter blockade on the location of the shell of the reward mountain within the space defined by the two independent variables, the **objective** opportunity cost of the reward, P_{obj} , and the **pulse** frequency, F_{pulse} . However, the core variables about which the inferences are to be drawn operate in spaces defined by the **subjective** opportunity cost, P_{sub} , and the induced frequency of **firing**, F_{firing} (Tab S2). As we explain below, the core \rightarrow shell function, f_F , which maps the pulse frequency into the evoked frequency of firing, contributes to the estimates of both location parameters of the shell. Its influence must be removed in order to isolate the effects of the drug on $F_{firing_{hm}}$ and P_{sub_e} , the location parameters of the core. In the depiction and interpretation of the fitted surfaces, we will distinguish between

- **primary** displacement of the shell of the reward mountain due to drug actions on the core components and
- **secondary** (additional) displacement of the shell due to the differential response of the frequency-following function (f_F) in the drug and vehicle conditions

In the following section, we explain how to remove the secondary displacements from the location-parameter estimates, thus isolating the primary displacements that are the focus of the study.

Correction of the location-parameter estimates for changes in frequency-following fidelity

Fitting the mountain surface to time-allocation data returns the two location parameters of the reward-mountain shell: P_{obj_e} , which positions the shell along the price axis, and $F_{pulse_{hm}}$, which positions the shell along the pulse-frequency axis. Whereas the value of $F_{pulse_{hm}}$ is independent of the value of P_{obj_e} , the reverse does not hold when frequency-following fidelity is imperfect. Changes in frequency-following fidelity alter the maximum reward intensity ($\hat{R}_{bsr_{max}}$), which contributes to the value of both P_{obj_e} and its subjective equivalent, P_{sub_e} . The portion of the shifts in P_{obj_e} due to changes in frequency-following fidelity must be removed in order to decouple estimates of that parameter from shifts along the pulse-frequency axis. Once P_{sub_e} has been suitably corrected, manipulations that act at, or beyond, the output of the reward-growth function (f_R , Fig S16) shift the mountain core uniquely along the price axis, whereas manipulations that alter the input to the reward-growth function shift the mountain core uniquely along the frequency axis [2, 3, 41].

Imperfect frequency-following fidelity can also alter the extent to which the surface of the mountain shell shifts along the pulse-frequency axis. This problem will arise if frequency-following fidelity differs in the drug and vehicle conditions. Correction is required in order to estimate the shift of fundamental interest, which is the displacement of the reward-growth function (a core component) along its firing-frequency axis.

The frequency-following function. The correction of the location-parameter estimates arises from the form of the frequency-following function (f_F). The form and parameters of this function for eICSS of the MFB were described by Solomon et al. [14]. In the absence of analogous data for oICSS, we have assumed the same functional form but have tuned the parameters to accommodate the power-frequency trade-off data reported here and results of prior studies [37–40].

In double-logarithmic coordinates, the frequency-following function (Fig S13) has the form of an inverted hockey stick, with a straight handle that transitions into a flat blade [6]. Pulse frequency is represented along the abscissa of this function, and firing frequency is represented along the ordinate. The induced firing frequency follows the pulse frequency perfectly over the portion of the handle that is truly straight, grows more slowly over a transition zone, and levels off over the blade portion. Thus, within the transition zone, a given drug-induced decrement in firing frequency (a core variable) will correspond to a larger decrement in pulse frequency (a shell variable). Drug-induced displacement of the reward-growth function (f_R) towards lower values of $F_{firing_{hm}}$ improves frequency-following fidelity (by moving the pulse frequency toward, or onto the straight “handle”). This causes the displacement of the shell to exceed, and thereby overestimate, the underlying displacement of the reward-growth function.

To correct our estimate of how far the reward-growth function has shifted along the pulse-frequency axis, we need to decouple its value from the maximum normalized reward intensity, $\hat{R}_{bsr_{max}}$. This is achieved by using the assumed frequency-following function (Eq S1) to estimate $F_{firing_{hm}}$ from $F_{pulse_{hm}}$.

We define the corrected estimate of the location parameter as follows:

$$F_{pulse_{hm}}^* = f_F(F_{pulse_{hm}}) \quad (S39)$$

where

$F_{pulse_{hm}}^*$ = the estimated value of $F_{firing_{hm}}$, which is the value $F_{pulse_{hm}}$ would have attained had frequency-following fidelity been perfect

Eq S39 states that $F_{pulse_{hm}}^*$ and $f_F(F_{pulse_{hm}})$ are one and the same. We will plot the value in question in the coordinate space of the shell. There, the pulse frequency, rather than the firing frequency, serves as the ordinate. Thus, in that context, we use $F_{pulse_{hm}}^*$ in lieu of $f_F(F_{pulse_{hm}})$ as our notation.

$f_F(F_{pulse_{hm}})$ (and thus $F_{pulse_{hm}}^*$) is a value along the ordinate of the frequency-following function (Fig S13), whereas $F_{pulse_{hm}}$ is a value along the abscissa. Given the form of the frequency-following function, $F_{pulse_{hm}}^* < F_{pulse_{hm}}$ once pulse frequency exceeds the capacity of the neurons to fire reliably to each and every pulse.

Several steps are required to correct the the parameter that locates the reward mountain along the price axis so that it too is decoupled from frequency-following fidelity. The step first is analogous to the estimation of $F_{pulse_{hm}}^*$ from $F_{pulse_{hm}}$: We use the subjective-price equation (Eq S6) to transform P_{obj_e} into its subjective equivalent, P_{sub_e} .

The next step is to correct P_{sub_e} for the effect of imperfect frequency-following fidelity. Eq S30 can be rearranged as follows:

$$P_{sub_e} = \left(\frac{p_{sub}}{[K_{ec} \times \hat{\phi}] \times [K_{aa} \times \hat{U}_e]} \times K_{rg} \right) \times \hat{R}_{bsr_{max}} \quad (S40)$$

Eq S40 reminds us that P_{sub_e} is proportional to the maximum normalized reward intensity that can be attained, $\hat{R}_{bsr_{max}}$. Eq S12 defines $\hat{R}_{bsr_{max}}$ in terms of the maximal attainable firing frequency, $F_{firing_{max}}$, the firing frequency corresponding to $F_{pulse_{hm}}$, and the reward-growth exponent, g . To estimate $F_{firing_{max}}$, we solve Eq S1 for a pulse frequency more than high enough to drive firing frequency to its maximum ($F_{pulse_{max}} = 1000$ pulses s^{-1}). We then use the resulting estimate of $\hat{R}_{bsr_{max}}$ to produce a revised estimate of P_{sub_e} :

$$P_{sub_e}^* = \frac{P_{sub_e}}{\hat{R}_{bsr_{max}}} \quad (S41)$$

where

$P_{sub_e}^*$ = estimated value that P_{sub_e} would have attained had frequency-following fidelity been perfect

Last, we transform $P_{sub_e}^*$ into its objective-price counterpart, $P_{obj_e}^*$ by passing $P_{sub_e}^*$ through the back-solution of the subjective-price equation [6]:

$$P_{obj_e}^* = P_{sub_{min}} + P_{sub_{bend}} \times \ln \left[-1 + e^{\left(\frac{P_{sub_e}^* - P_{sub_{min}}}{P_{sub_{bend}}} \right)} \right] \quad (S42)$$

$$\text{for } P_{sub_e}^* \geq P_{sub_0} \quad (S43)$$

The transformation of $P_{sub_e}^*$ into $P_{obj_e}^*$ is performed so that the corrected location-parameter estimate can be plotted in the space defined by the independent variables, $\{P_{obj}, F_{pulse}\}$. In this space, drug-induced shifts in the position of the reward mountain, corrected for changes in frequency-following fidelity, are depicted as

$$\left[\log_{10} \left(F_{pulse_{hm_{drug}}}^* \right) - \log_{10} \left(F_{pulse_{hm_{vehicle}}}^* \right) \right] \text{ and } \left[\log_{10} \left(P_{obj_{drug}}^* \right) - \log_{10} \left(P_{obj_{vehicle}}^* \right) \right].$$

Model fitting and selection

Model selection

The 12 candidate models fit to the data are described in Tab S3. The models differ in the total number of parameters as well as in the number of parameters free to vary across the vehicle and drug conditions. Models 2, 5, 8, and 11 are based on the six-parameter version of the reward-mountain model (Eq S35), whereas the remaining models are based on the seven-parameter version (Eq S38). The fits of all of the candidate models to the reward-mountain data from all seven rats converged successfully.

Table S3. The 12 candidate models fit to each dataset. Values of the “_free” parameters were free to differ between the vehicle and drug conditions, whereas a single value was fitted to the data from both conditions in the case of “_com” (common) parameters. Additional columns list the total number of “Common” and “Free” parameters along with their “Totals.” Models 2, 5, 8, and 11 are based on the six-parameter version of the reward-mountain model (Eq S35), whereas the remaining models are based on the seven-parameter version (Eq S38).

Num	a_free	g_free	CR_free	CR_com	Common	Free	Total
1	0	0	1	0	4	6	10
2	0	0	0	0	4	4	8
3	0	0	0	1	5	4	9
4	1	0	1	0	3	8	11
5	1	0	0	0	3	6	9
6	1	0	0	1	4	6	10
7	0	1	1	0	3	8	11
8	0	1	0	0	3	6	9
9	0	1	0	1	4	6	10
10	1	1	1	0	2	10	12
11	1	1	0	0	2	8	10
12	1	1	0	1	3	8	11

Tab S4 ranks the fits of the 12 candidate models for one rat (Bechr29) by their evidence ratios (the relative likelihood that a candidate model is true in comparison to the best-fitting model). Note that the residual sum of squares for the worst-fitting model (model 10) is slightly lower than in the case of the best-fitting model (model 2). This is not surprising given that the worst-fitting model comprises 12 parameters whereas the best-fitting model comprises only eight. The Akaike Information Criterion (AIC) [42] implements a trade-off between goodness of fit and simplicity. Thus, the AIC penalizes models with large number of parameters in comparison to simpler ones. On the basis of the AIC, model 10 is over 5×10^5 times less likely than model 2 and is ranked accordingly. Summary statistics for the best-fitting model for each rat are listed in Tab S5.

Table S4. Model-evaluation statistics for the fit of the 12 candidate models to the data from rat Bechr29. The models are described in Tab S3. “AIC” stands for the Akaike Information Criterion [42]. “Likelihood” refers to the ratio: $\left[\frac{\text{Likelihood}(\text{candidate model})}{\text{Likelihood}(\text{best-fitting model})} \right]$, which expresses the relative likelihood that a candidate model is true in comparison to the best-fitting model. The Evidence Ratio (“Ev_ratio”) is the inverse of the likelihood ratio. The total number of parameters is listed under “prms,” the residual sum of squares under “RSS,” the total sum of squares under “TSS,” and the adjusted R^2 under “Adj R^2 .”

Model	AIC	Likelihood	Ev_ratio	prms	RSS	TSS	Adj R^2
2	-1,922.58	1.000000	1.00	8	13.904	12,647.57	0.998886
8	-1,916.12	0.039636	25.23	9	13.897	12,647.57	0.998885
5	-1,915.86	0.034860	28.69	9	13.904	12,647.57	0.998884
3	-1,915.86	0.034846	28.70	9	13.904	12,647.57	0.998884
11	-1,909.46	0.001421	703.74	10	13.896	12,647.57	0.998883
9	-1,909.40	0.001377	726.48	10	13.897	12,647.57	0.998883
6	-1,909.14	0.001211	825.83	10	13.904	12,647.57	0.998882
1	-1,909.14	0.001210	826.19	10	13.904	12,647.57	0.998882
12	-1,902.74	0.000049	20,330.45	11	13.896	12,647.57	0.998881
7	-1,902.69	0.000048	20,826.76	11	13.897	12,647.57	0.998880
4	-1,902.42	0.000042	23,852.26	11	13.904	12,647.57	0.998880
10	-1,896.02	0.000002	585,023.15	12	13.895	12,647.57	0.998878

Table S5. Summary statistics for the model that provided the best fit (highest evidence ratio) to the data for each rat. The models in the “Model Num” column are defined in Tab S3. The residual sum of squares is listed in column “RSS,” the total sum of squares in column “TSS,” and the adjusted R^2 in column “Adj R^2 .”

Rat	Model Num	RSS	TSS	Adj R^2
Bechr14	2	23.23	14,413.79	0.9984
Bechr19	3	28.68	4,554.59	0.9936
Bechr21	11	19.52	14,177.29	0.9986
Bechr26	3	8.52	9,017.52	0.9990
Bechr27	12	4.55	11,171.77	0.9996
Bechr28	9	20.45	8,986.80	0.9977
Bechr29	2	13.90	12,647.57	0.9989

Different variants of the reward-mountain model provided the best fit to the data from different rats. Tab S6 shows the best-fitting model for all rats, as determined by the AIC-based evidence ratio. In four cases (Bechr14,19,26,29), the best-fitting model was one in which common values of the a and g parameters were fit to the data from both the vehicle and drug conditions; in the remaining three cases, the best-fitting model was one in which the values of a , g , or both were free to vary across the vehicle and drug conditions.

In four cases (Bechr19, 26,27,28), the C_r parameter was included in the best-fitting model, whereas in the three remaining cases, it was not. As Fig S15 illustrates, the C_r parameter will be advantageous when time allocation in low-payoff trials is higher along pulse-frequency sweeps than along price or radial sweeps. In no case was the value of this parameter free to vary across the vehicle and drug conditions in the best-fitting model. Thus, there is no evidence that the advantage conferred by addition of the C_r parameter was due to dopamine-transporter blockade.

Table S6. Best-fitting models for all rats. The models are described in Tab S3. Values of the “_free” parameters were free to differ between the vehicle and drug conditions, whereas a single value was fitted to the data from both conditions in the case of “_com” (common) parameters. Additional columns list the total number of common and free parameters and their totals

Rat	Model_num	a_free	g_free	CR_free	CR_com	Common	Free	Total
Bechr14	2	0	0	0	0	4	4	8
Bechr19	3	0	0	0	1	5	4	9
Bechr21	11	1	1	0	0	2	8	10
Bechr26	3	0	0	0	1	5	4	9
Bechr27	12	1	1	0	1	3	8	11
Bechr28	9	0	1	0	1	4	6	10
Bechr29	2	0	0	0	0	4	4	8

Fitted reward-mountain surfaces

380

Fig S17. Reward-mountain surfaces fit to the vehicle and drug data from rat Bechr14. The surfaces of the reward-mountain shell are shown in gray. The thick black line represents the contour mid-way between the minimal and maximal estimates of time allocation (the estimated altitudes of the valley floor and summit). Mean time-allocation values for the pulse frequency, price, and radial sweeps are denoted by red pyramids, blue squares, and green polyhedrons, respectively.

Fig S18. Reward-mountain surfaces fit to the vehicle and drug data from rat Bechr19. See caption for Fig S17.

Fig S19. Reward-mountain surfaces fit to the vehicle and drug data from rat Bechr21. See caption for Fig S17.

Fig S20. Reward-mountain surfaces fit to the vehicle and drug data from rat Bechr26. See caption for Fig S17.

Fig S21. Reward-mountain surfaces fit to the vehicle and drug data from rat Bechr27. See caption for Fig S17.

Fig S22. Reward-mountain surfaces fit to the vehicle and drug data from rat Bechr28. See caption for Fig S17.

The corresponding graph for rat Bechr29 is shown in Fig 6 in the main text along with the caption.

381

Contour and bar graphs

382

Fig S23. Contour graphs of the surfaces fit to the vehicle and drug data and bar graphs of the shifts in the location parameters for rat Bechr14. The values of the independent variables along frequency sweeps are designated by red triangles, along price sweeps by blue squares, and along radial sweeps by green circles. The values of the location parameters, $F_{pulse_{hm}}$ and P_{obj_e} , and are indicated by red horizontal lines with right-facing triangular end points and blue vertical lines with diamond end points, respectively. The shaded regions surrounding the lines denote 95% confidence intervals. The vehicle data are shown twice, once in the upper-left quadrant and once in the lower right. The dotted lines connecting the panels designate the shifts in the common-logarithmic values of the location parameters of the mountain, which are designated as $\{\Delta P_{obj_e}, \Delta F_{hm}\}$ and plotted in the bar graph in the upper-right panel. The dot-dash cyan lines superimposed on the bars show location-parameter estimates corrected for changes in frequency-following fidelity due to the displacement of the mountain along the pulse-frequency axis. (See section *Correction of the location-parameter estimates for changes in frequency-following fidelity*.) The 95% confidence intervals are shown in the bar graphs as vertical lines.

Fig S24. Contour and bar graphs for rat Bechr19. See caption for Fig S23.

Fig S25. Contour and bar graphs for rat Bechr21. See caption for Fig S23.

Fig S26. Contour and bar graphs for rat Bechr26. See caption for Fig S23.

Fig S27. Contour and bar graphs for rat Bechr27. See caption for Fig S23.

Fig S28. Contour and bar graphs for rat Bechr28. See caption for Fig S23.

The corresponding graphs for rat Bechr29 are shown in Fig 7 in the main text.

Location-parameter estimates

F_{hm} : Tab S7 shows the estimates of $F_{pulse_{hm}}$ (uncorrected) and $F_{pulse_{hm}}^*$ (corrected) for the drug and vehicle conditions. In six of seven cases, a lower pulse frequency sufficed to produce a reward of half-maximal intensity under the influence of dopamine-transporter blockade than in the vehicle condition.

The $F_{pulse_{hm}}$ estimates vary over more than a doubling range in both the drug and vehicle conditions. Particularly in the vehicle condition, the higher values fall within a range over which the assumed frequency-following function (??) rolls off, thus preventing the normalized reward-growth function from approaching a value of one at the highest pulse frequencies tested. This is why the uncorrected ($F_{pulse_{hm}}$) and corrected ($F_{pulse_{hm}}^*$) values in Tab S7 differ. For example, the ~ 31 pulses s^{-1} that produced a reward of half-maximal intensity in rat Bechr29 in the vehicle condition are estimated to have generated only ~ 26 firings s^{-1} .

Tab 2 in the main text shows the estimated drug-induced *shifts* in the location of the reward-mountain core along the frequency axis.

Eqs S13,S14 express the effect of the drug on the location parameter of the reward-growth function as a divisor: the more the drug boosts dopamine release, the lower the value of the location parameter for the drug condition and thus the farther the reward-growth function is shifted to the left. The rightmost column in Tab 2 lists the values of this divisor implied by the drug-induced shifts in the position of the reward mountain along the pulse-frequency axis. By analogy to Eq S14,

Table S7. The position of the reward mountain along the pulse-frequency axis. The $F_{pulse_{hm}}$ parameter sets the location of the shell of the reward-mountain along the pulse-frequency axis. If frequency-following fidelity were perfect, a pulse frequency of $F_{pulse_{hm}}$ would have induced an identical firing frequency in the optically activated dopamine neurons. Otherwise, the values corrected for imperfect frequency following, which are shown in the “ $F_{pulse_{hm}}^*$ ” columns, will be lower than the uncorrected values, which are shown in the “ $F_{pulse_{hm}}$ ” columns. The corrected values in the “ $F_{pulse_{hm}}^*$ ” columns are the estimated firing frequencies induced by the pulse frequencies in the “ $F_{pulse_{hm}}$ ” columns, derived from the frequency-following function described in section ???. The values listed in the “Veh” columns are from the fits to the data acquired in the vehicle condition, whereas the values in the “Drg” columns are from the fits to the data acquired under the influence of GBR-12909.

Rat	$F_{pulse_{hm}}$ Drg	$F_{pulse_{hm}}$ Veh	$F_{pulse_{hm}}^*$ Drg	$F_{pulse_{hm}}^*$ Veh
Bechr14	19.734	27.094	17.341	23.164
Bechr19	11.502	16.265	10.372	14.457
Bechr21	10.072	35.836	9.103	29.453
Bechr26	17.718	25.359	15.668	21.822
Bechr27	23.395	18.892	20.286	16.648
Bechr28	14.728	32.350	13.142	27.007
Bechr29	20.776	31.347	18.184	26.293

$$F_{pulse_{hm}drug}^* = \frac{F_{pulse_{hm}vehicle}^*}{K_{da_{drug}}} \quad (S44)$$

where

- $F_{pulse_{hm}drug}^*$ = location parameter of the reward-growth function for the drug condition
- $F_{pulse_{hm}vehicle}^*$ = location parameter of the reward-growth function for the vehicle condition
- $K_{da_{drug}}$ = proportional reduction in the value of the location parameter of the reward-growth function due to dopamine-transporter blockade

It follows that

$$K_{da_{drug}} = 10^{-\text{diff}}$$

where

$$\text{diff} = \log\left(F_{pulse_{hm}drug}^*\right) - \log\left(F_{pulse_{hm}vehicle}^*\right)$$

P_e: In previous work employing the reward-mountain model [3, 33, 43, 44], changes in the location of the fitted surface along the price axis have been attributed to variables acting at, or beyond, the **output** of the reward-growth function, whereas changes in the location of the fitted surface along the pulse-frequency axis have been attributed to variables acting at, or prior to, the **input** to the reward-growth function. The reward-mountain model treats these two sets of changes as independent, a postulate that is largely supported by empirical findings [1, 2, 13]. This interpretation is valid as long as the induced firing frequency can be driven high enough to maximize reward intensity. Tab S8 shows that this assumption does not hold in several of the datasets from the present study: The maximum normalized reward intensity, $\hat{R}_{bsr_{max}}$, is substantially less than one in these cases.

The deviation of $\hat{R}_{bsr_{max}}$ from one is generally greater in the vehicle data than in the drug data. In such cases, a

Table S8. Estimates of the maximum normalized reward intensities in the vehicle and drug conditions. The values in “ $\hat{R}_{bsr_{max}}$ Drg” and “ $\hat{R}_{bsr_{max}}$ Veh” columns are the maximum normalized reward intensities in the vehicle and drug conditions, respectively, based on the frequency-following function described in section ?? . The ratios of these values (Drg / Veh) are listed in the “Ratio” column, and the common logarithms of the ratios in the “log(Ratio)” column.

Rat	$\hat{R}_{bsr_{max}}$ Drg	$\hat{R}_{bsr_{max}}$ Veh	Ratio	log(Ratio)
Bechr14	0.986	0.958	1.029	0.013
Bechr19	1.000	0.999	1.001	0.000
Bechr21	1.000	0.849	1.179	0.071
Bechr26	0.998	0.989	1.009	0.004
Bechr27	0.964	0.997	0.967	-0.015
Bechr28	1.000	0.952	1.051	0.022
Bechr29	0.964	0.894	1.079	0.033

portion of the change in the value of the P_{obj_e} parameter is due to fact that the drug displaced the rising portion of the reward-mountain surface into a range of pulse frequencies over which the fidelity of frequency following is better than in the vehicle condition. That contribution to the change in the value of the P_{obj_e} parameter reflects mitigation at the *input* to the reward-growth function (Eq S11), thus undermining the independence of the changes in the parameters that locate the reward mountain along the price and frequency axes. That is why we computed corrected estimates ($P_{obj_e}^*$), as described in section *Correction of the location-parameter estimates for changes in frequency-following fidelity* thus compensating for the differences in the value of $\hat{R}_{bsr_{max}}$ across the vehicle and drug conditions. The estimates of P_{obj_e} and $P_{obj_e}^*$ are shown in Tab S9.

Table S9. The position of the reward mountain along the price axis. The P_{obj_e} parameter determines the position of the reward mountain along the price axis. When the maximum normalized reward intensity differs between the drug and vehicle conditions due to differences in frequency-following fidelity, the value of the P_{obj_e} parameter is affected. The values listed in the “ $P_{obj_e}^*$ ” columns have been corrected to remove this effect. They show the estimated value that the P_{obj_e} parameter would have attained had frequency-following fidelity been perfect.

Rat	P_{obj_e} Drg	P_{obj_e} Veh	$P_{obj_e}^*$ Drg	$P_{obj_e}^*$ Veh
Bechr14	9.572	6.127	9.414	6.252
Bechr19	11.755	5.369	11.658	5.363
Bechr21	25.102	13.662	25.096	15.933
Bechr26	14.051	11.625	14.347	11.843
Bechr27	22.434	19.475	23.066	19.592
Bechr28	75.392	51.225	76.366	56.401
Bechr29	10.041	6.295	10.480	7.375

Tab 3 in the main text lists the drug-induced shifts in the common logarithms of $P_{obj_e}^*$.

The drug-induced shifts in the location-parameter values are uncorrelated

Parameter values for best-fitting model for all rats

The values of the location parameters in these tables are uncorrected. Corrected values are listed in Tabs 2 and 3 in the main text.

Fig S29. Scatter plot of drug-induced shifts in the location parameters. The corrected estimates of displacement along the price and pulse-frequency axes are shown on the abscissa and ordinate, respectively. F_{hm}^* is shorthand for $F_{pulse_{hm}}^*$

Table S10. Parameter values from the best-fitting model for the data from Rat Bechr14. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	1.064	0.922	1.233	1.064	0.922	1.233
g	3.915	3.450	4.423	3.915	3.450	4.423
$Log_{10}(F_{hm})$	1.433	1.390	1.478	1.295	1.262	1.331
$Log_{10}(P_e)$	0.787	0.694	0.869	0.981	0.889	1.052
T_{max}	0.821	0.771	0.881	0.821	0.771	0.881
T_{min}	0.144	0.127	0.159	0.144	0.127	0.159

Table S11. Parameter values from the best-fitting model for the data from Rat Bechr19. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	1.840	1.729	1.972	1.840	1.729	1.972
C_r	0.214	0.191	0.239	0.214	0.191	0.239
g	5.862	3.837	13.690	5.862	3.837	13.690
$Log_{10}(F_{hm})$	1.211	1.172	1.247	1.061	0.996	1.100
$Log_{10}(P_e)$	0.730	0.699	0.762	1.070	1.051	1.089
T_{max}	0.796	0.777	0.816	0.796	0.777	0.816
T_{min}	0.005	0.000	0.009	0.005	0.000	0.009

Table S12. Parameter values from the best-fitting model for the data from Rat Bechr21. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	1.442	1.368	1.523	3.436	3.222	3.674
g	3.109	2.719	3.593	10.992	10.503	11.432
$Log_{10}(F_{hm})$	1.554	1.480	1.649	1.003	1.000	1.007
$Log_{10}(P_e)$	1.136	1.115	1.154	1.400	1.388	1.410
T_{max}	0.836	0.829	0.843	0.836	0.829	0.843
T_{min}	0.097	0.091	0.103	0.097	0.091	0.103

Table S13. Parameter values from the best-fitting model for the data from Rat Bechr26. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	1.998	1.887	2.106	1.998	1.887	2.106
C_r	0.071	0.062	0.080	0.071	0.062	0.080
g	5.233	4.788	5.804	5.233	4.788	5.804
$Log_{10}(F_{hm})$	1.404	1.384	1.427	1.248	1.232	1.268
$Log_{10}(P_e)$	1.048	1.080	1.148	1.134	1.164	
T_{max}	0.882	0.870	0.895	0.882	0.870	0.895
T_{min}	0.057	0.051	0.063	0.057	0.051	0.063

Table S14. Parameter values from the best-fitting model for the data from Rat Bechr27. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	1.454	1.381	1.525	2.015	1.924	2.135
C_r	0.037	0.035	0.039	0.037	0.035	0.039
g	5.050	4.561	5.684	3.524	3.360	3.699
$Log_{10}(F_{hm})$	1.276	1.248	1.307	1.369	1.342	1.394
$Log_{10}(P_e)$	1.289	1.272	1.303	1.351	1.337	1.364
T_{max}	0.888	0.881	0.895	0.888	0.881	0.895
T_{min}	0.009	0.000	0.028	0.009	0.000	0.028

Table S15. Parameter values from the best-fitting model for the data from Rat Bechr28. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	2.432	2.304	2.595	2.432	2.304	2.595
C_r	0.022	0.021	0.024	0.022	0.021	0.024
g	4.609	4.281	5.077	17.305	16.182	18.765
$Log_{10}(F_{hm})$	1.510	1.489	1.532	1.168	1.163	1.173
$Log_{10}(P_e)$	1.709	1.687	1.730	1.877	1.860	1.893
T_{max}	0.915	0.910	0.921	0.915	0.910	0.921
T_{min}	0.143	0.135	0.151	0.143	0.135	0.151

Table S16. Parameter values from the best-fitting model for the data from Rat Bechr29. Columns CB_{low} and CB_{high} list the upper and lower bounds of the 95.2% confidence intervals.

Parameter	Vehicle			Drug		
	Estimate	CB_{low}	CB_{high}	Estimate	CB_{low}	CB_{high}
a	2.312	2.172	2.462	2.312	2.172	2.462
g	3.161	2.980	3.354	3.161	2.980	3.354
$Log_{10}(F_{hm})$	1.496	1.476	1.518	1.318	1.304	1.335
$Log_{10}(P_e)$	0.799	0.783	0.814	1.002	0.990	1.015
T_{max}	0.896	0.883	0.910	0.896	0.883	0.910
T_{min}	0.112	0.107	0.118	0.112	0.107	0.118

The reward-growth function for oICSS

423

Fig S30. Changing the value of the input-scaling parameter, F_{hm} , shifts the mountain along the pulse-frequency axis. F_{hm} is shorthand for $F_{pulse_{hm}}$

Fig S31. Changing the value of the output-scaling parameter, K_{rg} , shifts the mountain along the price axis. P_e is shorthand for P_{obj_e}

Illustration of the correction for imperfect frequency-following fidelity

424

When the parameter that locates the reward mountain along the pulse-frequency axis ($F_{pulse_{hm}}$) falls within the range over which the induced firing frequency diverges substantially from the pulse frequency, then the magnitude of a drug-induced shift along the pulse-frequency axis is exaggerated, and a fictive shift is produced along the price axis. Fig S32 illustrates this effect and its removal by means of the correction procedure. The dot-dash cyan contour line shows the objective-price and pulse-frequency values that would have driven time allocation halfway between its minimal and maximal values had frequency-following fidelity been perfect. In contrast, the solid, wide, black contour line shows the equivalent objective-price and pulse-frequency values given the assumed frequency-following function. Note that the dot-dash cyan line deviates more from the solid, wide, black line in the simulated vehicle data in the upper left and lower right quadrants than in the simulated drug data in the lower-left quadrant. This is so because the uncorrected location-parameter value (vehicle: ~ 39 pulses s^{-1} ; drug: ~ 18 pulses s^{-1}) is much closer to the estimated maximum attainable firing frequency (51.67 firings s^{-1}) in the vehicle condition than in the drug condition. As a result, the simulated firing rate falls further below the pulse frequency in the vehicle condition than in the drug condition. The dot-dash lines superimposed on the bar graphs show the result of correcting the location-parameter shifts for this effect.

425
426
427
428
429
430
431
432
433
434
435
436
437
438

Fig S32. Correction for imperfect frequency-following fidelity. Simulated data, with $F_{pulse_{hm}}$ placed well within the region over which frequency-following fidelity falls off substantially. The simulated vehicle data are shown twice, once in the upper-left quadrant and once in the lower right. The dotted lines connecting the panels designate the shifts in the common-logarithmic values of the location parameters of the mountain, which are designated as $\{\Delta P_{obj_e}, \Delta F_{hm}\}$ and plotted in the bar graph in the upper-right panel. The dot-dash cyan lines superimposed on the bars show location-parameter estimates corrected for changes in frequency-following fidelity due to the displacement of the mountain along the pulse-frequency axis. F_{hm} is shorthand for $F_{pulse_{hm}}$

Comparison of logistic and power growth of reward intensity

439

Fig S33. The input-scaling parameter of the power reward-growth function locates the reward mountain along the price axis. Contour- and bar-graph representation of the simulated reward mountains produced by the magenta and green power-reward-growth functions (Eq 2) in the lower-left panel of Fig 10. In contrast to the effect of varying the value of the input-scaling parameter on the location of reward mountains based on logistic reward growth (Figs S30, S31), changing the value of the input-scaling parameter of the power-reward-growth function shifts the mountain along the price axis and not along the pulse-frequency axis.

Fig S34. The output-scaling parameter of the power reward-growth function also locates the reward mountain along the price axis. Contour- and bar-graph representation of the simulated reward mountains produced by the magenta and green power-reward-growth functions (Eq 2) in the lower-right panel of Fig 10. Changing the value of the output-scaling parameter of the power-reward-growth function shifts the mountain along the price axis just like the effect of changing the value of the input-scaling parameter shown in Fig S33.

Fig S35. Contour graphs of reward mountains simulated by the convergence model.

Dopamine-transporter blockade shifts the simulated reward mountain (almost) exclusively along the pulse-frequency axis, as in the behavioral data. The simulated MFB drive on the dopamine neurons is equivalent to an optical pulse frequency of 40 pulses s^{-1} .

Fig S36. Contour graphs of reward mountains simulated by the convergence model given very strong MFB input.

The simulated MFB drive on the dopamine neurons is now equivalent to an optical pulse frequency of 80 pulses s^{-1} .

Training in preparation for measurement of the reward mountain

Fig S37. Graphical summary of the experimental procedure. A: TH::Cre +/- rats received bilateral VTA injections of an AAV5 virus bearing a Cre-dependent, ChR2-YFP transcript. Optical fibers were bilaterally aimed at the VTA. B: Rats were trained to hold down a lever for a specified cumulative amount of time to deliver trains of optical stimulation to the VTA. The red curve represents the proportion of trial time the rat spent working for the optical reward as the optical pulse frequency (the reward-strength variable) was systematically manipulated. The blue curve shows the proportion of trial time the rat spent working for a maximal optical reward as the cumulative amount of time required to harvest the reward (“price”) was manipulated systematically. The green curves show proportion of trial time the rat spent working for the optical reward as the strength and price of the reward were simultaneously manipulated. C: The reward-mountain model was fit independently to the data from each rat following injections of GBR-12909 or vehicle. Within subject comparisons were performed.

Contour lines: the trade-off between pulse frequency and price to hold time allocation constant

Contour graphs provide a compact summary of the reward-mountain surface in a format that facilitates visualization of the direction(s) in which the mountain has been shifted by a manipulation such as administration of a drug. The changes in the values of the location parameters become visually apparent in this format.

Here, we derive the equation for the contour lines, thus updating an earlier derivation [14] in which it had been assumed that the higher pulse frequencies tested drive reward intensity to its maximum attainable value ($\hat{R}_{max} \rightarrow 1$). This will indeed be so if the pulse frequencies in question are substantially lower than the frequency-following limit in the directly stimulated neurons. That assumption was usually justified in previous eICSS studies in which the reward mountain was measured [2, 3, 6, 13, 14, 33, 43, 44]. Highly excitable MFB neurons served as the directly activated substrate for the rewarding effect in those studies. In contrast, midbrain dopamine neurons are directly activated substrate in the current study. Not only do these neurons have more limited frequency-following abilities than their MFB counterparts [38, 40], their activation is due to optical excitation of a relatively slow opsin [36] rather than to electrical excitation of voltage-sensitive membrane channels. As we show below, it is likely that the highest pulse frequencies employed in the present study did not always succeed in driving reward intensity to its maximum, particularly in the vehicle condition. To accommodate such cases, we now generalize the previously published expression for the contour lines [14].

We begin by reformatting Eq 33 from the main text as follows:

$$\vec{\hat{R}}_{cont} = \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \quad (S45)$$

where

a = price-sensitivity exponent

$\vec{P}_{obj_{cont}}$ = vector of prices for which $\hat{T} = \hat{T}_{cont}$ when \hat{R} is a corresponding element of \hat{R}_{cont} (S46)

P_{obj_e} = objective price at which $T = T_{mid}$ when $\hat{R} = \hat{R}_{max}$

$f_P(P_{obj_{cont}})$ = subjective equivalent of $P_{obj_{cont}}$

$f_P(P_{obj_e})$ = subjective price at which $T = T_{mid}$ when $\hat{R} = \hat{R}_{max}$

$\vec{\hat{R}}_{cont}$ = vector of normalized reward intensities for which $\hat{T} = \hat{T}_{cont}$ when P_{obj} is a corresponding element of $P_{obj_{cont}}$; $0 \leq \hat{R} \leq 1$

\hat{R}_{max} = maximum normalized reward intensity

\hat{T}_{cont} = time allocation represented by the contour line

\hat{T}_{cont} = normalized time allocation represented by the contour line;

$\hat{T}_{cont} = (T_{cont} - T_{min}) \div (T_{max} - T_{min})$; $0 \leq \hat{T}_{cont} \leq 1$;

Substituting for \hat{R} from Eq 12 in the main text, we obtain:

460

$$\frac{f_F(\vec{F}_{pulse_{cont}})^g}{f_F(\vec{F}_{pulse_{cont}})^g + f_F(F_{pulse_{hm}})^g} = \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \quad (S47)$$

We now multiply both sides by $f_F(\vec{F}_{pulse_{cont}})^g + f_F(F_{pulse_{hm}})^g$, yielding

$$f_F(\vec{F}_{pulse_{cont}})^g = \left[f_F(\vec{F}_{pulse_{cont}})^g + f_F(F_{pulse_{hm}})^g \right] \times \left\{ \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \right\} \quad (S48)$$

and we then expand the right side to yield:

$$f_F(\vec{F}_{pulse_{cont}})^g = \left\{ f_F(\vec{F}_{pulse_{cont}})^g \times \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \right\} + \left\{ f_F(F_{pulse_{hm}})^g \times \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \right\} \quad (S49)$$

The terms that include $f_F(\vec{F}_{pulse_{cont}})^g$ are collected on the left side

$$f_F(\vec{F}_{pulse_{cont}})^g - \left\{ f_F(\vec{F}_{pulse_{cont}})^g \times \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \right\} =$$

$$f_F(\vec{F}_{pulse_{hm}})^g \times \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \quad (S50)$$

and the left side is factored to yield:

$$f_F(\vec{F}_{pulse_{cont}})^g \times \left(1 - \left\{ \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \right\} \right) =$$

$$f_F(\vec{F}_{pulse_{hm}})^g \times \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left(\frac{f_P(\vec{P}_{obj_{cont}})}{f_P(P_{obj_e})} \right) \right] \quad (S51)$$

Re-arranging the terms, we obtain:

$$\frac{f_F(\vec{F}_{pulse_{cont}})^g}{f_F(\vec{F}_{pulse_{hm}})^g} = \frac{\left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left\{ \frac{P_{sub}(\vec{P}_{obj_{cont}})}{P_{sub}(P_{obj_e})} \right\} \right]}{1 - \left\{ \left(\frac{\hat{T}_{cont}}{1 - \hat{T}_{cont}} \right)^{\frac{1}{a}} \times \left[\hat{R}_{max} \times \left\{ \frac{P_{sub}(\vec{P}_{obj_{cont}})}{P_{sub}(P_{obj_e})} \right\} \right] \right\}} \quad (S52)$$

The numerator and denominator of the right side are now multiplied by $f_P(P_{obj_e}) \times \left(\frac{1 - \hat{T}_{cont}}{\hat{T}_{cont}} \right)^{\frac{1}{a}}$ to yield

$$\frac{f_F(\vec{F}_{pulse_{cont}})^g}{f_F(\vec{F}_{pulse_{hm}})^g} = \frac{\hat{R}_{max} \times f_P(\vec{P}_{obj_{cont}})}{\left[\left(\frac{1 - \hat{T}_{cont}}{\hat{T}_{cont}} \right)^{\frac{1}{a}} \times f_P(P_{obj_e}) \right] - \left(\hat{R}_{max} \times f_P(\vec{P}_{obj_{cont}}) \right)} \quad (S53)$$

The contour graphs will be plotted in double logarithmic coordinates. In that space, Eq S53 becomes

$$\log_{10} [f_F(\vec{F}_{pulse_{cont}})] - \log_{10} (f_F(F_{pulse_{hm}})) = \left(\frac{1}{g} \right) \times \log_{10} \left\{ \frac{\hat{R}_{max} \times f_P(\vec{P}_{obj_{cont}})}{\left[\left(\frac{1-\hat{T}_{cont}}{\hat{T}_{cont}} \right)^{\frac{1}{a}} \times f_P(P_{obj_e}) \right] - [\hat{R}_{max} \times f_P(\vec{P}_{obj_{cont}})]} \right\} \quad (S54)$$

When time allocation falls halfway between T_{min} and T_{max} ($T = T_{mid}$; $\hat{T} = 0.5$), Eq S54 reduces to:

$$\log_{10} [f_F(\vec{F}_{pulse_{\hat{T}=0.5}})] - \log_{10} [f_F(F_{pulse_{hm}})] = \left(\frac{1}{g} \right) \times \log_{10} \left\{ \frac{\hat{R}_{max} \times f_P(\vec{P}_{obj_{\hat{T}=0.5}})}{f_P(P_{obj_e}) - [\hat{R}_{max} \times f_P(\vec{P}_{obj_{\hat{T}=0.5}})]} \right\} \quad (S55)$$

Expanding the right side, we obtain

$$\log_{10} [f_F(\vec{F}_{pulse_{T=0.5}})] - \log_{10} [f_F(F_{pulse_{hm}})] = \left(\frac{1}{g} \right) \times \left\{ \log_{10} \left\langle \hat{R}_{max} \times f_P(\vec{P}_{obj_{T=0.5}}) \right\rangle - \log_{10} \left\langle f_P(P_{obj_e}) - [\hat{R}_{max} \times f_P(\vec{P}_{obj_{T=0.5}})] \right\rangle \right\} \quad (S56)$$

When $\hat{R}_{max} = 1$, Eq S56 reduces to

$$\log_{10} [f_F(\vec{F}_{pulse_{T=0.5}})] - \log_{10} [f_F(F_{pulse_{hm}})] = \left(\frac{1}{g} \right) \times \left\{ \log_{10} [f_P(\vec{P}_{obj_{T=0.5}})] - \log_{10} [f_P(P_{obj_e}) - f_P(\vec{P}_{obj_{T=0.5}})] \right\} \quad (S57)$$

References

1. Arvanitogiannis A, Shizgal P. The reinforcement mountain: allocation of behavior as a function of the rate and intensity of rewarding brain stimulation. Behavioral Neuroscience. 2008;122(5):1126–1138. doi:10.1037/a0012679.
2. Breton YA, Mullett A, Conover K, Shizgal P. Validation and extension of the reward-mountain model. Frontiers in Behavioral Neuroscience. 2013;7:125. doi:10.3389/fnbeh.2013.00125.
3. Hernandez G, Breton YA, Conover K, Shizgal P. At what stage of neural processing does cocaine act to boost pursuit of rewards? PLoS ONE. 2010;5(11). doi:10.1371/journal.pone.0015081.
4. Gibson J. The Theory of Affordances. In: Shaw R, Bransford J, editors. Perceiving, Acting, and Knowing: Toward an Ecological Psychology. Hillsdale, N.J.: Lawrence Erlbaum Associates; 1977. p. 67–82.
5. McDowell JJ. On the classic and modern theories of matching. Journal of the Experimental Analysis of Behavior. 2005;84(1):111–127. doi:10.1901/jeab.2005.59-04.

6. Solomon RB, Trujillo-Pisanty I, Conover K, Shizgal P. Psychophysical Inference of Frequency-Following Fidelity in the Neural Substrate for Brain Stimulation Reward. *Behavioural Brain Research*. 2015;292:327–341. doi:10.1016/j.bbr.2015.06.008.
7. Edmonds DE, Gallistel CR. Reward versus performance in self-stimulation: electrode-specific effects of alpha-methyl-p-tyrosine on reward in the rat. *Journal of Comparative and Physiological Psychology*. 1977;91(5):962–974.
8. Miliaressis E, Rompré PP, Laviolette P, Philippe L, Coulombe D. The curve-shift paradigm in self-stimulation. *Physiology & behavior*. 1986;37(1):85–91.
9. Wise RA. Action of Drugs of Abuse on Brain Reward Systems. *Pharmacology, biochemistry, and behavior*. 1980;13 Suppl 1:213–223.
10. Gallistel CR. Spatial and Temporal Summation in the Neural Circuit Subserving Brain Stimulation Reward. In: Wauquier A, Rolls ET, editors. *Brain-Stimulation Reward: A Collection of Papers Prepared for the First International Conference on Brain-Stimulation Reward at Janssen Pharmaceutica, Beerse, Belgium on April 21-24, 1975*. Oxford: North-Holland; 1976. p. 97–99.
11. Gallistel CR. Self-Stimulation in the Rat: Quantitative Characteristics of the Reward Pathway. *Journal of Comparative and Physiological Psychology*. 1978;92(6):977–998.
12. Matthews G. Neural substrate for brain stimulation reward in the rat: cathodal and anodal strength-duration properties. *Journal of Comparative and Physiological Psychology*. 1977;91(4):858–874.
13. Breton YA, Conover K, Shizgal P. The effect of probability discounting on reward seeking: a three-dimensional perspective. *Frontiers in Behavioral Neuroscience*. 2014;8:284. doi:10.3389/fnbeh.2014.00284.
14. Solomon RB, Conover K, Shizgal P. Valuation of Opportunity Costs by Rats Working for Rewarding Electrical Brain Stimulation. *PLOS ONE*. 2017;12(8):e0182120. doi:10.1371/journal.pone.0182120.
15. Niyogi RK, Breton YA, Solomon RB, Conover K, Shizgal P, Dayan P. Optimal indolence: a normative microscopic approach to work and leisure. *Journal of The Royal Society Interface*. 2013;11(91):20130969–20130969. doi:10.1016/S0166-4328(02)00282-6.
16. Petry N, Heyman G. Rat toys, reinforcers, and response strength: An examination of the Re parameter in Herrnstein's equation. *Behavioural Processes*. 1997;39:39–52.
17. Gallistel C, Shizgal P, Yeomans J. A Portrait of the Substrate for Self-Stimulation. *Psychol Rev*. 1981;88(3):228–273.
18. Gallistel CR, Leon M. Measuring the Subjective Magnitude of Brain Stimulation Reward by Titration with Rate of Reward. *Behavioral Neuroscience*. 1991;105(6):913–925.
19. Leon M, Gallistel CR. The Function Relating the Subjective Magnitude of Brain Stimulation Reward to Stimulation Strength Varies with Site of Stimulation. *Behavioural Brain Research*. 1992;52(2):183–193.
20. Simmons JM, Gallistel CR. Saturation of Subjective Reward Magnitude as a Function of Current and Pulse Frequency. *Behavioral Neuroscience*. 1994;108(1):151–160.
21. Sonnenschein B, Conover K, Shizgal P. Growth of Brain Stimulation Reward as a Function of Duration and Stimulation Strength. *Behavioral Neuroscience*. 2003;117(5):978–994. doi:http://psycnet.apa.org/journals/bne/117/5/978.
22. Killeen P. The Matching Law. *Journal of the Experimental Analysis of Behavior*. 1972;17(3):489–495.
23. Gallistel CR, Gibbon J. Time, rate, and conditioning. *Psychological Review*. 2000;107(2):289–344.
24. Kiss I, Shizgal P. Compound action potentials recorded in the ventral tegmental area, substantia nigra, and periaqueductal gray following rewarding stimulation of the lateral hypothalamus in the rat. *Psychobiology*. 1990;18(2):205–214.

25. Rompré PP, Shizgal P. Electrophysiological characteristics of neurons in forebrain regions implicated in self-stimulation of the medial forebrain bundle in the rat. *Brain research*. 1986;364(2):338–349. 521
522
26. Shizgal P, Schindler D, Rompré PP. Forebrain neurons driven by rewarding stimulation of the medial forebrain bundle in the rat: comparison of psychophysical and electrophysiological estimates of refractory periods. *Brain research*. 1989;499(2):234–248. 523
524
525
27. Heyman GM, Monaghan MM. Reinforcer magnitude (sucrose concentration) and the matching law theory of response strength. *J Exp Anal Behav*. 1994;61(3):505–516. 526
527
28. Heyman GM, Beer B. A new approach for evaluating the behavioral effects of antipsychotic drugs. *Trends in Pharmacological Sciences*. 1987;8(10):388–393. 528
529
29. Heyman G. How drugs affect cells and reinforcement affects behavior: formal analogies. In: Commons ML, Church RM, Stellar JR, Wagner AR, editors. *Biological determinants of reinforcement*. vol. 7 of *Quantitative analysis of behavior*. New York, NY: Lawrence Erlbaum Associates; 1988. p. 157–182. 530
531
532
30. Herrnstein R. On the Law of Effect. *J Exp Anal Behav*. 1970;13(2):243–266. 533
31. Herrnstein R. Formal Properties of the Matching Law. *J Exp Anal Behav*. 1974;21(1):159–164. 534
32. Breton YA, Marcus JC, Shizgal P. *Rattus Psychologicus*: construction of preferences by self-stimulating rats. *Behavioural Brain Research*. 2009;202(1):77–91. doi:10.1016/j.bbr.2009.03.019. 535
536
33. Hernandez G, Trujillo-Pisanty I, Cossette MP, Conover K, Shizgal P. Role of dopamine tone in the pursuit of brain stimulation reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2012;32(32):11032–11041. doi:10.1523/JNEUROSCI.1051-12.2012. 537
538
539
34. Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science (New York, NY)*. 2009;324(5930):1080–1084. 540
541
542
35. Witten IB, Steinberg EE, Davidson TJ, Yizhar O, Ramakrishnan C, Stuber GD, et al. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*. 2011;72(5):721–733. doi:10.1016/j.neuron.2011.10.028. 543
544
545
36. Klapoetke NC, Murata Y, Kim SS, Pulver SR, Birdsey-Benson A, Cho YK, et al. Independent optical excitation of distinct neural populations. *Nature Methods*. 2014;11(3):338–346. doi:10.1038/nmeth.2836. 546
547
37. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*. 2012; p. 1–6. doi:10.1038/nature10754. 548
549
38. Ilango A, Kesner AJ, Broker CJ, Wang DV, Ikemoto S. Phasic excitation of ventral tegmental dopamine neurons potentiates the initiation of conditioned approach behavior: parametric and reinforcement-schedule analyses. *Frontiers in Behavioral Neuroscience*. 2014;8:155. doi:10.3389/fnbeh.2014.00155. 550
551
552
39. Covey DP, Cheer JF. Accumbal Dopamine Release Tracks the Expectation of Dopamine Neuron-Mediated Reinforcement. *Cell Reports*. 2019;27(2):481–490.e3. doi:10.1016/j.celrep.2019.03.055. 553
554
40. Lohani S, Martig AK, Deisseroth K, Witten IB, Moghaddam B. Dopamine Modulation of Prefrontal Cortex Activity Is Manifold and Operates at Multiple Temporal and Spatial Scales. *Cell Reports*. 2019;27(1):99–114.e6. doi:10.1016/j.celrep.2019.03.012. 555
556
557
41. Shizgal P, Conover K, Trujillo-Pisanty I. Video Introducing the Mountain Model; 2012. 558
42. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974;19(6):716–723. doi:10.1109/TAC.1974.1100705. 559
560
43. Trujillo-Pisanty I, Conover K, Shizgal P. A new view of the effect of dopamine receptor antagonism on operant performance for rewarding brain stimulation in the rat. *Psychopharmacology*. 2014;231(7):1351–1364. doi:10.1007/s00213-013-3328-x. 561
562
563

44. Trujillo-Pisanty I, Hernandez G, Moreau-Debord I, Cossette MP, Conover K, Cheer JF, et al. Cannabinoid receptor blockade reduces the opportunity cost at which rats maintain operant performance for rewarding brain stimulation. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2011;31(14):5426–5435. doi:10.1523/JNEUROSCI.0079-11.2011.

564
565
566
567