City Research Online

City St George's, University of London

Citation: Rigoli, F., Chew, B., Dayan, P. & Dolan, R.J. (2016). The Dopaminergic Midbrain Mediates an Effect of Average Reward on Pavlovian Vigor. Journal of Cognitive Neuroscience, 28(9), pp. 1303-1317. doi: 10.1162/jocn_a_00972

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/16679/

Link to published version: https://doi.org/10.1162/jocn_a_00972

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

The Dopaminergic Midbrain Mediates an Effect of Average Reward on Pavlovian Vigor

Francesco Rigoli, Benjamin Chew, Peter Dayan, and Raymond J. Dolan

Abstract

■ Dopamine plays a key role in motivation. Phasic dopamine response reflects a reinforcement prediction error (RPE), whereas tonic dopamine activity is postulated to represent an average reward that mediates motivational vigor. However, it has been hard to find evidence concerning the neural encoding of average reward that is uncorrupted by influences of RPEs. We circumvented this difficulty in a novel visual search task where we measured participants' button pressing vigor in a context where information (underlying an RPE) about future average reward was provided well before the average reward itself. Despite no instrumental consequence, participants' pressing force increased for greater current average reward, consistent with a form of Pavlovian effect on motivational vigor. We recorded participants' brain activity during task performance with fMRI. Greater average reward was associated with enhanced activity in dopaminergic midbrain to a degree that correlated with the relationship between average reward and pressing vigor. Interestingly, an opposite pattern was observed in subgenual cingulate cortex, a region implicated in negative mood and motivational inhibition. These findings highlight a crucial role for dopaminergic midbrain in representing aspects of average reward and motivational vigor. ■

INTRODUCTION

The ventral tegmental area/substantia nigra (VTA/SN) region of midbrain is the major source of ascending brain dopaminergic neuromodulation. Substantial evidence indicates that rapid bursts in VTA/SN neuronal firing rates correlate with expression of an appetitive reinforcement prediction error (RPE; D'Ardenne, McClure, Nystrom, & Cohen, 2008; Tobler, Fiorillo, & Schultz, 2005; Schultz, Dayan, & Montague, 1997). It has been suggested that tonic dopaminergic activity in VTA/SN also plays a key role in instrumental aspects of motivation by representing the average rate of reward (Skvortsova, Palminteri, & Pessiglione, 2014; Meyniel, Sergent, Rigoux, Daunizeau, & Pessiglione, 2013; Niv, Daw, Joel, & Dayan, 2007), which quantifies an opportunity cost of sloth and balances the price of alacrity.

In a recent test of this model, we showed that RTs (as one index of motor vigor) decrease proportionally to the average experienced reward (Guitart-Masip, Beierholm, Dolan, Duzel, & Dayan, 2011). A link to dopamine activity was established in a subsequent finding where the impact of average reward on vigor was enhanced by administration of levodopa, a dopaminergic precursor that increases the availability of this neurotransmitter (Beierholm et al., 2013). Nevertheless, several fundamental questions remain including whether activity in dopaminergic rich VTA/SN is modulated by average reward. This question has been problematic to address because of the need to disentangle an effect of average reward from effects linked to the expression of RPEs. Moreover, whether a putative representation of average reward in VTA/SN is connected with expressions of motor vigor remains to be tested.

To investigate both questions, we used fMRI recording brain activity in healthy human participants while they performed a novel computer-based task (Figure 1A) that required a right/left button press corresponding to the position of a visual target stimulus, presented together with distractors. On each trial, participants received a performance-independent baseline monetary reward, as well as a fixed (£3) reward for a correct response. Within blocks, the baseline reward was fixed, but across blocks it varied over three levels (£1, £6, £11). To ensure incentive compatibility, at the end of the task, a random trial was selected and the corresponding total reward was added to an initial payment of £17.

Given that baseline reward remained fixed within blocks, this variable can be linked to average reward as it is the key determinant of long run rate of reinforcement. Crucially, our design allowed us to isolate the impact of this baseline reward from any influence of an RPE. We accomplished this segregation by signaling, at the start of each block n, the baseline reward of the subsequent block $n + 1$, an experimental manipulation expected to generate an RPE at this time point. This implies a temporal lag between information about baseline reward (underlying RPE) and the same baseline reward being available (associated with average reward). Therefore,

University College London

© 2016 Massachusetts Institute of Technology

Journal of Cognitive Neuroscience 28:9, pp. 1303–1317
doi:10.1162/jocn_a_00972

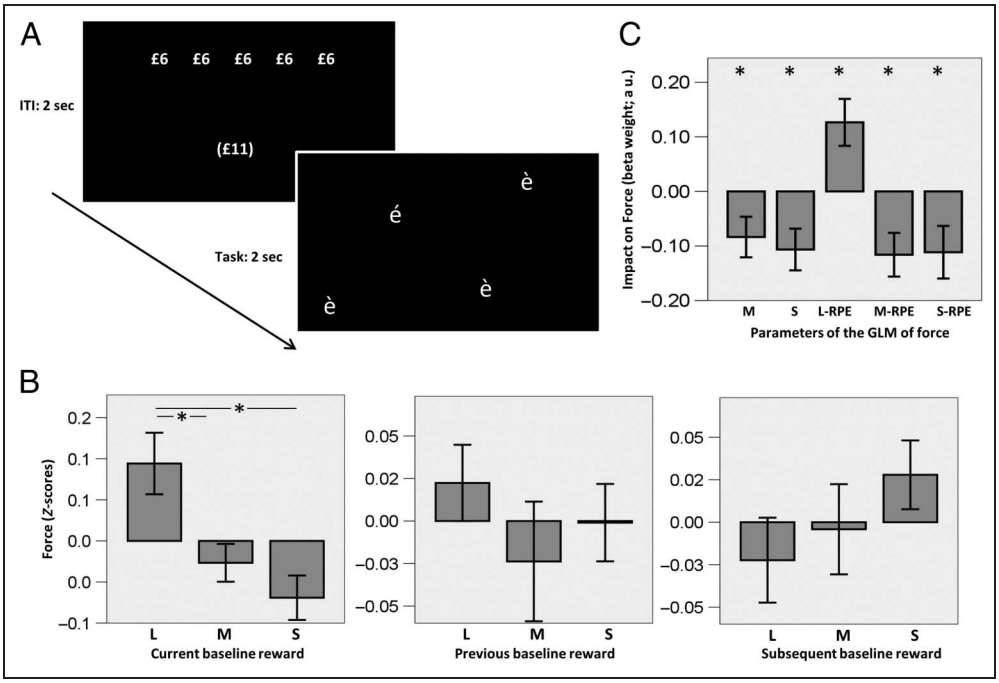

Figure 1. (A) Experimental paradigm. Before commencement of each trial, an information panel was shown for 2 sec displaying (i) on the top of the screen a row of monetary amounts corresponding to the baseline reward (i.e., independent from performance) of the current block n (the number of monetary amounts displayed corresponds to the number of trials remaining in the current blocks) and (ii) on the bottom of the screen a monetary amount in brackets corresponding to the baseline reward available on the subsequent block $n + 1$. Next the target (*é*) and three distractors (*è*) were presented. Participants were required to press, within a 2-sec window, a left/right button corresponding to the side of the screen on which the target is displayed (left or right). Stimuli remained on the screen for 2 sec and then were followed by a new information panel. For correct responses, a £3 reward fixed across blocks was added to the baseline reward. At the end of the experiment, one trial was randomly selected and paid out. (B) Effect of large (L), medium (M), and small (S) baseline reward on force for (from left to right) baseline reward at current ($F(2, 34) = 5.4$, $p = .009$), previous ($F(2, 34) = 0.507$, $p = .607$), and subsequent block ($F(2, 34) = 0.815$, $p = .451$). Asterisks indicate significant differences when comparing large minus medium ($t(17) = 2.15$, $p = .046$) and large minus small baseline reward ($t(17) = 2.67$, $p = .016$) for current block (no difference was found between medium and small baseline reward for current block ($t(17) = 1.29$, $p = .213$)). (C) Beta weights estimated with the GLM of force implementing the influence of the baseline reward of the current block n plus the influence (on the first trial of bock n alone) of the baseline reward relative to the subsequent block $n + 1$. From left to right, the parameters plotted represent: (M) the effect of medium minus large baseline reward for current block n, (S) the effect of small minus large baseline reward for current block n, (L-RPE) the effect on the first trial of a block n of large baseline reward for subsequent block $n + 1$, (M-RPE) the effect on the first trial of a block n of medium minus large baseline reward for subsequent block $n + 1$, (S-RPE) the effect on the first trial of a block n of small minus large baseline reward for subsequent block $n + 1$. Asterisks indicate that all parameters are significantly different from zero (M: $t(17) = -2.249$, $p = .038$; S: $t(17) = -2.786$, $p = .013$; L-RPE: $t(17) = 2.942$, $p = .009$; M-RPE: $t(17) = 2.896$, $p = .010$; S-RPE: $t(17) = 2.312$, $p = .034$).

we could dissociate the impact of the baseline reward of the current block n (associated with an average reward) from the impact of the associated reward information provided at the start of the previous block $n - 1$, the time at which it would have generated an RPE.

As an index of motor vigor, on each trial we measured the force (see Methods) participants exerted when pressing the button. One hypothesis is that larger baseline reward would enhance vigor. This hypothesis could fit with a form of Pavlovian effect, because there is no instrumental consequence of acting more vigorously. This is because, by design, the reward amount dependent on performance was fixed across conditions. However, we also considered the possibility that participants might press harder with smaller baseline reward, because reward amount dependent on performance (fixed across conditions) might be rescaled relative to the baseline reward, reflecting a form of endowment effect (Kahneman, Knetsch, & Thaler, 1991). This predicts that, for instance, the performance-dependent £3 would be perceived as subjectively more valuable in a condition associated to £1 than £11 baseline reward. In addition, we used computational modeling to test whether force production was influenced by the information about the baseline reward of block $n + 1$ that was provided at the start of block n and led to an RPE. The presence of an influence at this time point would support the idea that participants paid attention to the RPE-related information and hence that our task manipulation was effective.

Dissociating the influence of a current baseline reward from an RPE allowed us to study the relationship between relatively higher or lower average available reward and signals in regions implicated in respectively appetitive or aversive motivation. For appetitive motivation, we focused on the VTA/SN and ventral striatum (Bartra, McGuire, & Kable, 2013; D'Ardenne et al., 2008; Tobler et al., 2005; O'Doherty et al., 2004; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Schultz et al., 1997). For aversive motivation, we considered the subgenual cingulate cortex (sGC), a region implicated in the expression of inhibition seen in negative mood states (Rauch & Drevets, 2009). For example, sGC is activated by negative

mood (Kohn et al., 2014; Rauch & Drevets, 2009; Phan et al., 2005; Mayberg et al., 1999; George et al., 1995), whereas activity in this region distinguishes depressed patients and healthy controls (Drevets, Savitz, & Trimble, 2008; Drevets et al., 1997). In addition, deep brain stimulation to sGC has been reported as being efficacious in refractory depression (Berlim, McGirr, Van den Eynde, Fleck, & Giacobbe, 2014; Mayberg et al., 2005). We also studied the amygdala, as it is widely implicated in coordinating Pavlovian conditioned responses in both appetitive and aversive contexts (Talmi, Seymour, Dayan, & Dolan, 2008; Fanselow & Gale, 2003; Davis, 1992).

METHODS

Participants

Twenty healthy right-handed adults participated in the experiment. Two participants were excluded from analyses because they repeatedly stopped doing the task inside the scanner and did not complete the entire task. Thus, the experimental sample included 18 participants (12 women, age = 19–34 years, mean age = 26 years). The study was approved by the University College of London research ethics committee.

Experimental Paradigm and Procedure

Inside the MRI scanner, participants performed a computer-based task lasting 32 min (Figure 1A). On each trial, a target (corresponding to the letter *é*) and three distractors (corresponding to the letter *è*) appeared simultaneously on the screen, and the four stimuli were shown in a randomized position with two of them appearing on each side of the screen. For each trial, participants received a baseline monetary reward varying across blocks on three levels (£1, £6, £11) that was independent of performance plus a £3 reward (fixed across blocks) if they correctly pressed a right/left button on a keypad (using the middle/index finger of the right hand) corresponding to the position of the target within 2 sec. Trials with equal baseline reward were arranged in blocks (each including eight trials) ordered pseudorandomly. During the intertrial interval, an information panel was presented for 2 sec showing (i) the number of trials remaining in the current block n represented as a row of equal monetary amounts displayed on the top of the screen and (ii) the baseline reward of the subsequent block $n + 1$ represented by a monetary amount displayed in the bottom of the screen in brackets. After the information panel, the target and distractors were presented and remained on the screen for 2 sec independently from RT, followed either by a new information panel or by an error feedback appearing for 1 sec when participants pressed the wrong button or did not press at all. At the end of the experiment, one outcome was randomly selected among all those received in the entire task and added to an initial participation payment of £17. We recorded the force exerted by participants during button pressing using a purpose-built two-button box. Pressure in units proportional to Pascals was measured on a continuous scale.

Participants were tested at the Wellcome Trust Centre for Neuroimaging at the University College London. Before scanning, participants provided informed consent and were fully instructed as to the task contingencies and rules about the payment. They were not told that the force with which they pressed was recorded. Next, they familiarized with the task outside the scanner for up to 100 unpaid trials. Inside the scanner, participants performed the task in two separate sessions each including 30 blocks. After scanning, participants were debriefed and informed about their total remuneration.

Behavioral Methods

We analyzed the data about pressing force by comparing the predictions made by different computational models of the data. We considered several hypotheses about the mechanisms relating reward contingency and force production. Below we formalize each hypothesis using a computational model and outline and test the predictions that each model makes. Crucially, the different models make specific predictions that allow us to infer which model explains better the connection between reward contingency and force production in the real data. We characterized two influences: one based on a latent variable we call reward value V, which updates with every reward experience, and the other based on a reward prediction error signal RPE, which applies just to the first trial of a block, when new information was provided.

In all models, at every trial t belonging to block n, a reward $R(n)$ is collected depending on the baseline reward associated with block n (for simplicity, we discarded error trials and thus omitted the performance-dependent £3, given that this amount was constant across trials and performance was almost perfect for all participants—i.e., 95%). We considered four different models of how this determines the reward value V.

Baseline Model Learning (BM_{LEA}): At every trial t belonging to block n, a reward equivalent to $R(n)$ is collected and used to update a representation of reward value V_1 according to a delta rule characterized by a learning rate α:

$$V_1(t) = V_1(t-1) + \alpha(R(n) - V_1(t-1))$$

Baseline Model Previous (BM_{PRE}): At every trial t belonging to block n, the reward value $V_2(t)$ corresponds to the baseline reward of the previous block $n - 1$:

$$V_2(t) = R(n-1)$$

Baseline Model Current (BM_{CUR}): At every trial t belonging to block n, the reward value $V_3(t)$ corresponds

to the baseline reward of the current block n (note that this model corresponds to BM_{LEA} with a learning rate α equal to one):

$$V_3(t) = R(n)$$

Baseline Model Subsequent (BM_{SUB}): At every trial t belonging to block n, the reward value $V_4(t)$ corresponds to the baseline reward of the subsequent block $n + 1$:

$$V_4(t) = R(n + 1)$$

We also considered the influence on force of the information provided just before the first trial of a block n about the baseline reward of the subsequent block $n + 1$.

We tested the possibility that this novel information would produce an RPE which in turn affects force production in the first trial of block n. In general, an RPE occurs when novel information is used to update a prior expectation of reward. The models considered here differ with respect to which variables are considered as novel reward information and prior reward expectation.

RPE Model Subsequent/Current ($RPEM_{SUB/CUR}$): The novel reward information provided at trial t (which is the first trial of block n) corresponds to the baseline reward of the subsequent block $n + 1$, and the prior reward expectation corresponds to the baseline reward of the current block n. Thus, the RPE at the first trial of a block is

$$RPE_1(t) = R(n + 1) - R(n)$$

RPE Model Subsequent/Previous ($RPEM_{SUB/PRE}$): The novel reward information provided at trial t (which is the first trial of block n) corresponds to the baseline reward of the subsequent block $n + 1$, and the prior reward expectation corresponds to the baseline reward of the previous block $n - 1$. Thus, the RPE at the first trial of a block is

$$RPE_2(t) = R(n + 1) - R(n - 1)$$

RPE Model Subsequent/Average ($RPEM_{SUB/AVE}$): The novel reward information provided at trial t (which is the first trial of block n) corresponds to the baseline reward of the subsequent block $n + 1$, and the prior reward expectation corresponds to the average baseline reward (motivated by the fact that baseline rewards were pseudorandomized and participants were informed about this during instructions). Thus, the RPE at the first trial of a block is

$$RPE_3(t) = R(n + 1) - mean(R)$$

RPE Model Current/Average ($RPEM_{CUR/AVE}$): We also considered the possibility that participants started paying attention to the current baseline reward only at the start of the current block. In this account, the novel reward information provided at trial t (which is the first trial of block n) corresponds to the baseline reward of the current block n, and the prior reward expectation corresponds to the average baseline reward (motivated by the fact that baseline rewards were pseudorandomized). Thus, the RPE at the first trial of a block would be

$$RPE_4(t) = R(n) - mean(R)$$

In total, we assumed that the force F exerted in button-pressing was represented by

$$F(t) = \beta_0 + \beta_v V(t) + \omega\beta_{RPE} RPE(t) + \varepsilon$$

where ω is equal to 1 and 0 for first and nonfirst trials of blocks, respectively, ε represents a noise parameter, and βs represent linear weight parameters.

Different models made different predictions with respect to our behavioral analyses (see Results), allowing us to assess which model fits better with the empirical results from the analyses. Models were compared based on whether they predicted the statistical effects that emerged from our data. We did not perform a trial-by-trial model fit and comparison as this would have been redundant. Note that a possibility is that participants used a mixed strategy involving the use of a mixture of models. For instance, a model integrating BM_{LEA} and BM_{SUB} would compute the reward value $V_{1,4}$ as follows:

$$V_{1,4}(t) = \rho(V_1(t-1) + \alpha(V_1(t-1) - R(n)) + (1-\rho)R(n + 1)$$

in which ρ corresponds to a weighting parameter. Our analyses allowed us to consider the possibility of a mixed strategy given that, if participants used more than one model, we would expect results compatible with all these models, manifested in the data as a combination of the associated effects.

Imaging Methods

The task was programmed with the Cogent toolbox (Wellcome Trust Centre for Neuroimaging, London) in Matlab (The MathWorks, Natick, MA). Visual stimuli were back-projected onto a translucent screen positioned behind the bore of the magnet and viewed via an angled mirror. BOLD contrast functional images were acquired with echo-planar T2*-weighted imaging using a Siemens (Berlin, Germany) Trio 3-T MR system with a 32-channel head coil. To maximize amount of data in our ROIs, a partial volume of the ventral part of the brain was recorded. Each image volume consisted of 25 interleaved 3-mm-thick sagittal slices (in-plane resolution = 3 × 3 mm; time to echo = 30 msec; repetition time = 1.75 sec). The first six volumes acquired were discarded to allow for T1 equilibration effects. T1-weighted structural images were acquired at a 1 × 1 × 1 mm resolution. fMRI data were analyzed using Statistical Parametric Mapping version 8 (Wellcome Trust Centre for Neuroimaging). Data preprocessing included spatial realignment, unwarping using individual field maps, slice timing correction, normalization, and smoothing. Specifically, functional

volumes were realigned to the mean volume, the first slice was used as reference for slice timing correction, and volumes were spatially normalized to the standard Montreal Neurological Institute (MNI) template with a $3 \times 3 \times 3$ voxel size and were smoothed with 8-mm Gaussian kernel. High-pass filtering with a cutoff of 256 sec (chosen because the design involved relatively long blocks—i.e., 30 sec) and AR(1) model were applied.

The main general linear model (GLM) included a canonical hemodynamic response function and three box car function regressors associated with the different baseline rewards of the current block n (£1, £6, £11) with durations corresponding to block lengths. Three stick function regressors associated with the baseline reward of the subsequent block $n + 1$ were also included at start of block n, plus a stick function regressor indicating when an error response occurred. Note that regressors were largely uncorrelated (maximum cos(θ) was around 0.2 across regressors; see Figure S1) due to the temporal gap between regressors associated with the current block n and the regressors associated with the subsequent block $n + 1$. This GLM was also used in a psychophysiological interaction (PPI) analysis to probe interregional coupling changes as a function of baseline reward condition of the current block n. A second GLM was estimated to test the effect of pressing force and included a box car function regressor with durations corresponding to block lengths modulated by the average pressing force exerted in each block. Again, three stick function regressors associated with the baseline reward of the subsequent block $n + 1$ were also included at the start of block n, plus a stick function regressor indicating error responses. This GLM was used to investigate separately the relationship of brain activity with force and baseline reward, given their behavioral correlation. Participants' respiration and heart rate signals were recorded and, together with estimated motion, were included as regressors of no interest in all GLMs. For each GLM, contrasts of interest were computed participant by participant and used for second-level one-sample t tests and regressions across participants.

Statistical testing was based on an ROI approach, motivated by prior findings in relation to appetitive and aversive forms of motivation. For VTA/SN, the ROI was manually defined using the software MRIcro and the mean structural image, using a method similar to that of Guitart-Masip, Fuentemilla, et al. (2011). Other ROIs were defined as 6-mm spheres centered on coordinates extracted from previous studies (for ventral striatum, Bartra et al., 2013; for amygdala, De Martino, Kumaran, Seymour, & Dolan, 2006; for sGC, Davey, Harrison, Yücel, & Allen, 2012). For hypothesis testing on ROIs, a small volume correction (SVC) with $p < .05$ family wise error (FWE) was applied. For exploratory purposes, we also looked at other areas using $p < .05$ FWE-corrected with respect to the partial volume recorded. However, no activation was found in any region using these criteria.

RESULTS

Behavior

Across participants, the average percentage of correct responses was 95% (in the range 87–99%). We considered potential effects of current baseline reward on two behavioral measures: RTs and force. We found no effect of current baseline reward on z-scored RTs ($F(2, 34) = 1.28, p = .291$; two-tailed $p < .05$ is used as significance criterion for behavioral statistical tests). This may be explained by the nature of the task used here, which is different from those adopted in other studies in which RTs did reflect motivational vigor (e.g., Guitart-Masip, Beierholm, et al., 2011). In particular, a visual search-type task requires a speed–accuracy trade-off that may have limited any effect on RTs. For instance, higher baseline reward might slow down RTs in some participants because of an attempt to improve accuracy. For this reason, we hypothesized that the force exerted during button pressing would be better suited in this task as an expression of motivational vigor. Consistently, the force of pressing (Z-scored for all behavioral and neural analyses separately for each finger, index, or middle) was affected by baseline reward (Figure 1B; $F(2, 34) = 5.4, p = .009$; main results on force are presented in Table 1), reflecting an enhanced response for a large baseline reward condition compared with both medium ($t(17) = 2.15, p = .046$) and small reward ($t(17) = 2.67, p = .016$) conditions. There was no significant difference between medium and small baseline rewards ($t(17) = 1.29, p = .213$). We interpret this effect of large baseline reward as reflecting a Pavlovian influence in so far as reward dependent on performance was fixed across the different baseline reward conditions; hence, exertion of an excess force during a large baseline reward condition had no instrumental consequence. We observed no relationship between previous or subsequent baseline reward on pressing force (Figure 1B; previous reward: $F(2, 34) = 0.507, p = .607$; subsequent reward: $F(2, 34) = 0.815, p = .451$).

Behavior: Computational Modeling

We probed the behavioral data further by comparing the predictions arising from different computational models of force production (described in Methods). Synthetic data were generated by simulating the models (see Figures 2 and 3). Following the behavioral results reported above, the models assume that the reward collected at each trial is equal to 1 during blocks with large baseline reward and equal to 0 during blocks with both medium and small baseline reward. Note that the models do not have free parameters except the BM_{LEA} for which we used $V_1(0) = 0$ and $\alpha = 0.2$ for the simulation. Four thousand trials were run organized in blocks each including eight trials with equal baseline reward; blocks were ordered randomly. On the basis of the simulated data, we outlined and tested the predictions made by each model. Note that, qualitatively, predictions of BM_{LEA} (shown in

Table 1. Behavioral Results Relative to the Main Analyses of Pressing Force

Effect on Force	*Independent Variable*	*Contrast*	*Statistic*	*p*
Effect on all trials of blocks	Baseline reward of current block n	F contrast	$F(2, 34) = 5.4$	.009*
	Baseline reward of previous block $n - 1$	F contrast	$F(2, 34) = 0.507$	.451
	Baseline reward of subsequent block $n + 1$	F contrast	$F(2, 34) = 0.815$	.607
	Baseline reward of current block n	Large minus Medium	$t(17) = 2.15$	.046*
		Large minus Small	$t(17) = 2.67$	.016*
		Medium minus Small	$t(17) = 1.29$	.213
Additional effect on the first trial of blocks	Baseline reward of subsequent block $n + 1$	Large minus Medium	$t(17) = 2.896$	.010*
		Large minus Small	$t(17) = 2.312$	.034*
		Medium minus Small	$t(17) = -0.092$	.928
	Baseline reward of current block n	Large minus Medium	$t(17) = -0.432$	.672
		Large minus Small	$t(17) = -0.210$	.836
		Medium minus Small	$t(17) = 0.262$	.796
	Baseline reward of previous block $n - 1$	Large minus Medium	$t(17) = 0.784$	.445
		Large minus Small	$t(17) = -0.533$	.601
		Medium minus Small	$t(17) = -1.224$	.238

The statistical tests are performed over parameters extracted from different GLMs. To obtain consistency within the table, note that some t statistics of the section "Additional Effect on the First Trial of Blocks" have flipped sign compared with the same analyses reported in the main text. Significant p values are marked with asterisks.

Figure 2) are independent of the parameters chosen (except for α equal to 1, where BM_{LEA} reduces to BM_{CUR}). Crucially, different models make different predictions with respect to these behavioral analyses, allowing us to assess which model fits better with the empirical results emerged from the analyses.

Analysis 1. Effect of Baseline Reward of the Current, Previous, and Subsequent Block

As reported above, we observed an effect on force of baseline reward of current ($F(2, 34) = 5.4$, $p = .009$) but not previous ($F(2, 34) = 0.507$, $p = .607$) nor subsequent ($F(2, 34) = 0.815$, $p = .451$) block. The effect was driven by an increase in force with large baseline reward compared with both medium ($t(17) = 2.15$, $p = .046$) and small baseline reward ($t(17) = 2.67$, $p = .016$) conditions, with no difference between medium and small baseline reward ($t(17) = 1.29$, $p = .213$).

Data simulated by the different models relative to this analysis are reported in Figure 2. These results are compatible with BM_{CUR} (that predicts an effect of baseline reward at current block alone) and are incompatible with BM_{PRE} (that predicts an effect of baseline reward at previous but not current block) and BM_{SUB} (that predicts an effect of baseline reward at subsequent but not current block). In relation with BM_{LEA}, these results are consistent with a large learning rate α only (because BM_{LEA} can be reduced to BM_{CUR} if the learning rate α is equal to one).

Analysis 2. Trial-by-Trial Impact of Reward

The previous analysis leaves open the question of whether in our data there was any effect of reward history, as in BM_{LEA}, or not, as in BM_{CUR}. From our simulated data, BM_{LEA} alone predicts that force should decrease over the course of blocks characterized by small baseline reward (and medium baseline reward, given that these two conditions were indistinguishable in the previous analysis). Also, BM_{LEA} alone predicts that force should increase along blocks characterized by large baseline reward. For each baseline reward condition separately, we correlated the force with the trial number within block (i.e., the first trial was assigned a value of 1, the second trial a value of 2, and so on) and found no correlation (large reward: $t(17) = -1.37$, $p = .190$; medium reward: $t(17) = -0.999$, $p = .332$; small reward: $t(17) = -0.382$, $p = .708$), contrary to predictions of BM_{LEA} and consistent with predictions of BM_{CUR}.

Analysis 3. Effect of Information on the Subsequent Baseline Reward

To test for an effect of the information provided just before the first trial of block n about the reward that will be

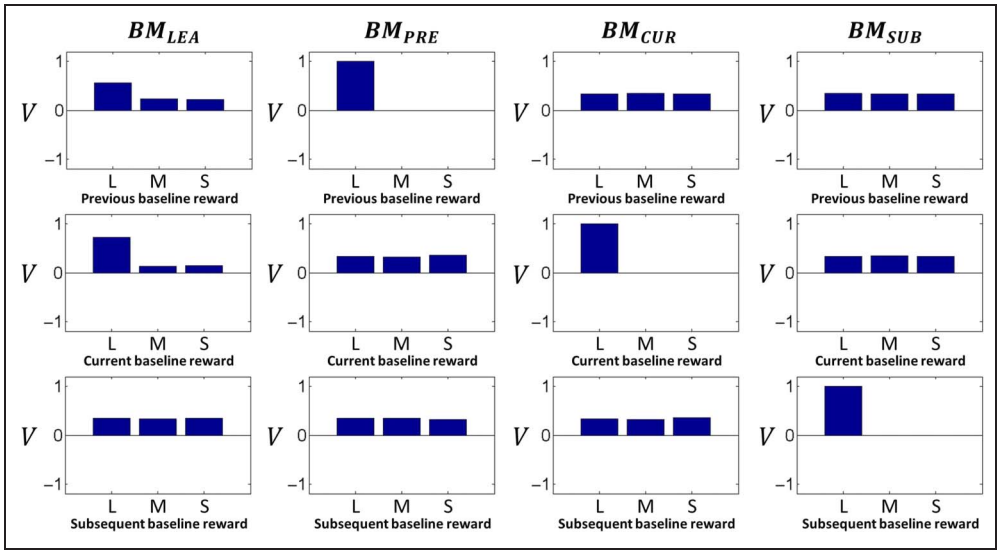

Figure 2. Simulated data generated with the different models of force (separated in columns) outlined in the main text. On each trial, models employ specific algorithms to compute the reward value *V*. In the figure, for each model, the average *V* is reported relative to the true baseline reward at the previous (first row), current (second row), and subsequent (third row) block. In the models used in this simulation, the reward collected at each trial *t* belonging to block *n* is *R*(*n*) and is equal to 1 for large baseline reward and equal to 0 for both medium and small reward. Our behavioral analyses fit better with predictions made by BM_{CUR}, which predicts an effect of the current baseline reward alone (see Figure 1B).

provided in the subsequent block $n + 1$, we tested the impact of the reward of the previous, current, and subsequent baseline reward on the first trials of blocks. Simulated data of the different models relative to this analysis are reported in Figure 3. Models $RPEM_{SUB/CUR}$, $RPEM_{SUB/PRE}$, and $RPEM_{SUB/AVE}$, but not $RPEM_{CUR/AVE}$, predict an increased force when the information about the reward of the subsequent block $n + 1$ signals a large baseline reward.

We tested this prediction using a GLM of force that included as regressors an intercept parameter, the force at previous trial (as nuisance regressor), and five binary characteristic function regressors reporting (i) medium baseline reward for current block n, (ii) small baseline reward for current block n, (iii) whether a trial was the first of block n, (iv) whether a trial was the first of block n and the subsequent block $n + 1$ was associated with medium baseline reward, and (v) whether a trial was the first of block n and the subsequent block $n + 1$ was associated with small baseline reward. Second-level t tests on the parameters (Figure 1C) confirmed that large baseline reward in the current block enhanced force compared with medium (one-sample t test on the binary regressor (i): $t(17) = -2.249$, $p = .038$) and small baseline reward (one-sample t test on the binary regressor (ii): $t(17) = -2.786$, $p = .013$) with no difference between medium and small baseline reward (two-sample t test of the binary regressor (i) minus (ii): $t(17) = 1.104$, $p = .285$). In addition, when testing for effects on the first trial of block n dependent on information about the baseline reward of

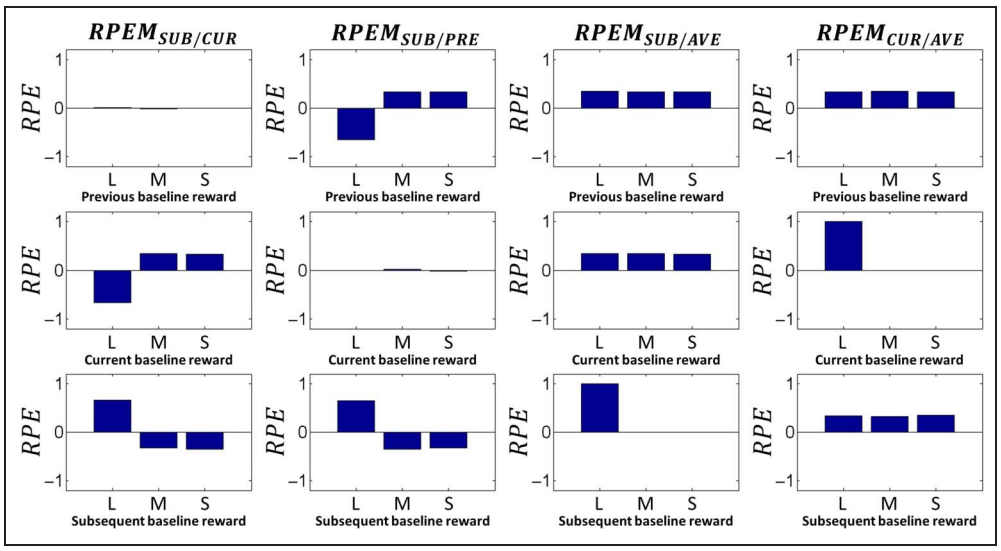

Figure 3. Simulated data generated with the different models (separated in columns) considering the impact of new reward information (provided just before the first trial of block *n* about the baseline reward of the subsequent block $n + 1$) on the force exerted on the very first trial of block *n*. On each first trial of blocks, models employ specific algorithms to compute the reward prediction error RPE. In the figure, for each model the average RPE is reported relative to the true baseline reward at the previous (first row), current (second row), and subsequent (third row) block. In the models used in this simulation, the reward collected at each trial *t* belonging to block *n* is *R*(*n*) and is equal to 1 for large baseline reward and equal to 0 for both medium and small reward. Our behavioral analyses fit better with predictions made by $RPEM_{SUB/AVE}$, which predicts an effect of the subsequent baseline reward alone (see Figure 1C).

the subsequent block $n + 1$, we observed that force increased for a large subsequent baseline reward compared with nonfirst trials (one-sample t test on the binary regressor (iii): $t(17) = 2.942, p = .009$) and decreased with medium compared with large (one-sample t test on the binary regressor (iv): $t(17) = -2.896, p = .010$) and small compared with large subsequent baseline rewards (one-sample t test on the binary regressor (v): $t(17) = -2.312$, $p = .034$), with no difference between the medium and small subsequent baseline reward (two-sample t test of the binary regressor (iv) minus (v): $t(17) = -0.092$, $p = .928$). These results show an increased force on the first trial of block n when the information about the reward of the subsequent block $n + 1$ signals a large baseline reward (compared with when it signals a medium and small baseline reward) and are consistent with $RPEM_{SUB/CUR}$, $RPEM_{SUB/PRE}$, and $RPEM_{SUB/AVE}$, but not $RPEM_{CUR/AVE}$.

$RPEM_{SUB/CUR}$ alone predicts decreased force in the first trials of blocks when the current block is associated with large baseline reward. We tested this prediction with a GLM of pressing force equal to the one described above except that now regressor (iv) indicates whether a trial was the first of block n and the current block n was associated with medium baseline reward and regressor (v) indicates whether a trial was the first of block n and the current block n was associated with small baseline reward. When testing for effects on the first trial of blocks dependent on the baseline reward of the current block, we found no difference between large and medium (one-sample t test on the binary regressor (iv): $t(17) = 0.432$, $p = .672$), large and small (one-sample t test on the binary regressor (v): $t(17) = 0.210, p = .836$), and medium and small baseline reward (two-sample t test of the binary regressor (iv) minus (v): $t(17) = 0.262, p = .796$).

$RPEM_{SUB/PRE}$ alone predicts a decreased force in first trials of blocks when the previous block is associated with large baseline reward. We tested this prediction with a GLM of pressing force equal to the one described above except that now regressor (iv) indicates whether a trial was the first of block n and the previous block $n - 1$ was associated with medium baseline reward and regressor (v) indicates whether a trial was the first of block n and the previous block $n - 1$ was associated with small baseline reward. When testing for effects on the first trial of blocks dependent on the baseline reward of the previous block, we found no difference between large and medium (one-sample t test on the binary regressor (iv): $t(17) = -0.784, p = .445$), large and small (one-sample t test on the binary regressor (v): $t(17) = 0.533, p = .601$), and medium and small baseline reward (two-sample t test of the binary regressor (iv) minus (v): $t(17) = -1.224, p = .238$).

Altogether, these results are consistent with $RPEM_{SUB/AVE}$. According to this, information delivered at the start of block n concerning the baseline reward of a subsequent block $n + 1$ exerts an influence on force. That is, motor vigor was boosted when a large baseline reward was signaled for the subsequent block $n + 1$, compared with when medium and small baseline rewards were signaled. In addition, this effect was not affected by the baseline reward of previous or current block. This is consistent with the possibility that, because each of the three baseline rewards had the same chance of occurrence in every block (based on the fact that order of baseline rewards was pseudorandomized), at the start of a block n participants had the same expectancy (i.e., independent of previous and current baseline reward) about the baseline reward for the subsequent block $n + 1$. Therefore, signaling that a subsequent block $n + 1$ has a large baseline reward leads to a positive RPE and greater force on the first trial of block n; signaling that the subsequent block $n + 1$ has a small or medium baseline reward leads to a negative RPE and smaller force.

In summary, these data confirm that participant responses were influenced by information about the baseline reward available for block $n + 1$. This dissociation allowed us to segregate the effects of this form of information (underlying generation of an RPE) from the effects associated with the corresponding baseline reward, because such information was provided temporarily prior to the baseline reward becoming available.

Imaging

While participants performed the visual search task, we used fMRI to measure BOLD activation in brain areas of a priori interest by virtue of their link to valuation and motivation. We estimated parameters of a GLM including three boxcar function regressors associated with the different baseline reward conditions of the current block (small, medium, and large reward) with durations corresponding to block length. The model also included three RPE stick function regressors at the start of block n associated with the signaling of the different baseline reward conditions on the subsequent block $n + 1$. This implementation of RPE regressors followed from our behavioral results (see also Methods). Crucially, following convolution with the hemodynamic response function, the boxcar regressors associated with the current baseline reward were uncorrelated with their corresponding stick function RPE regressors, as the latter occurred about 32 sec before the former.

Comparing the current block during large baseline reward against medium and small baseline reward (a contrast based on the behavioral results), we observed increased activation in left VTA/SN (Figure 4A; −11, −22, −10; $Z = 3.07, p = .014$ SVC; in MNI coordinates space; neural results are reported in Table 2) and right ventral striatum (right: 17, 8, −5; $Z = 2.81, p = .042$ SVC). For the opposite contrast, we observed increased activation in the current block for medium and small compared with large baseline reward in bilateral sGC (Figure 4B; right: 7, 21, −10; $Z = 3.59, p = .005$ SVC; left = −8,

31, −15; $Z = 4.36$, $p < .001$ SVC) and left amygdala (Figure 5A; −21, 3, −25; $Z = 2.80$, $p = .044$ SVC). In support of our behavioral observation of an apparent equivalence between medium and small baseline reward, we observed no differential activation effect in any ROIs when comparing these two conditions ($p > .05$ SVC).

A main goal of our experiment was to examine the relationships between average reward, vigor (as measured by the pressing force), and activity in dopaminergic VTA/SN (Niv et al., 2007). On the basis of our finding that baseline reward affects both force and activity in VTA/SN, we considered three distinct hypotheses: (a) baseline reward directly impacts both VTA/SN and force, with no further association between VTA/SN and force; (b) baseline reward directly impacts both VTA/SN and force, and VTA/SN has a separate, additional, association with force; and (c) baseline reward impacts on activity in VTA/SN, and this structure is in turn associated with force.

Note that all three hypotheses imply that activation in VTA/SN would correlate with force. To test this prediction, we estimated a GLM that included a regressor at block start, with duration corresponding to block length, modulated by the average force exerted in that block (thus capturing the relationship between neural activation with tonic force rather than with trial-by-trial variation in force) plus three regressors encoding the information (provided at the start of the current block n) about the baseline reward of the subsequent block $n + 1$ (as in the first GLM of brain neural data). In this GLM, we observed that left VTA/SN activity correlated with average force (−8, −15, −15; $Z = 2.64$, $p = .032$ SVC) whereas right sGC activity was inversely correlated with force (5, 28, −5, $Z = 2.86$, $p = .027$ SVC). No other ROI showed a significant positive or negative correlation.

Hypotheses (b) and (c), but not (a), predict that the individual beta weights relating baseline reward and VTA/SN activity should correlate with the beta weight relating baseline reward and force. To test this, we correlated the behavioral parameter encoding the difference in force for large on the one hand minus medium and small current baseline reward on the other, with the parameter encoding a difference in brain activity in these conditions. Specifically, the behavioral parameter corresponded to the negative of the sum of parameter (iv) and (v) of the first GLM described above in the subsection "Analysis 3. Effect of Information on the Subsequent Baseline Reward." Note that a positive behavioral parameter indicates that large baseline reward increases vigor compared with medium and small baseline reward. We found a positive correlation between the beta weights relating baseline reward and VTA/SN activity with the beta weight relating baseline reward and force in left VTA/SN (Figure 4D; −8, −17, −13; $Z = 3.22$, $p = .011$ SVC) and an inverse correlation in left sGC with a trend level effect in right sGC (left: −8, 31, −15; $Z = 3.29$, $p = .024$ SVC; right: 10, 26, −15; $Z = 2.65$, $p = .060$ SVC).

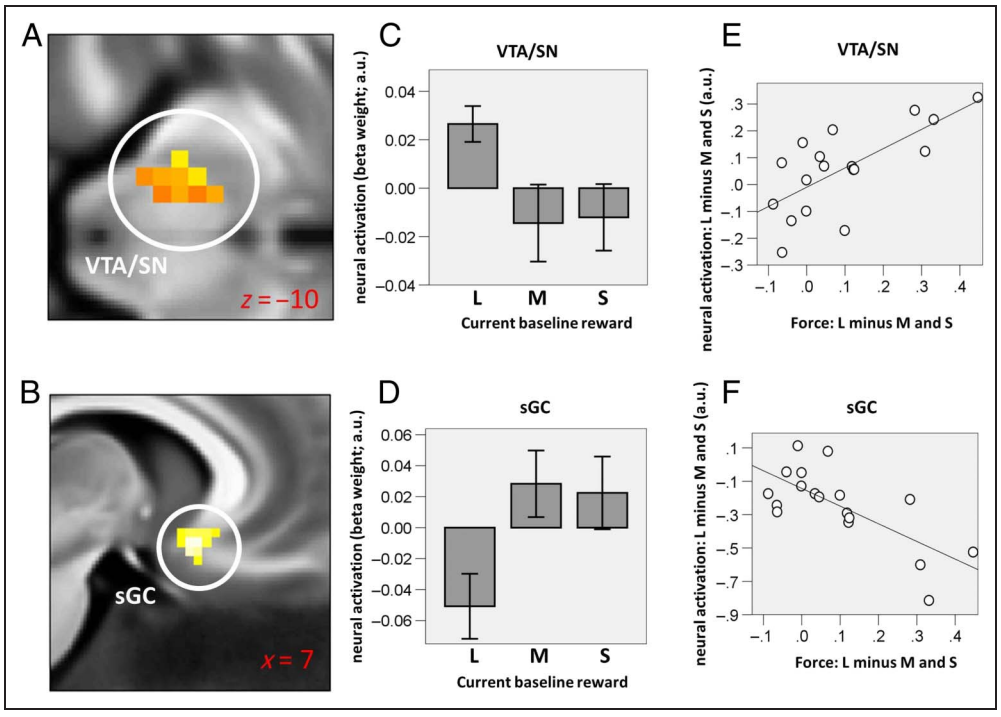

Figure 4. (A) Activation in left VTA/SN for the contrast large minus medium and small baseline reward at the current block (−11, −22, −10; $Z = 3.07$, $p = .014$ SVC; for this and following analyses, the statistic relative to the peak activation voxel within a ROI is small volume corrected with a FWE of $p < .05$, see also Methods; further we used the MNI coordinates space). (B) Activation in sGC for the contrast medium and small minus large baseline reward at the current block (right: 7, 21, −10; $Z = 3.59$, $p = .005$ SVC; left = −8, 31, −15; $Z = 4.36$, $p < .001$ SVC). (C) Beta weights for large (L), medium (M), and small (S) baseline reward at the current block for peak activation voxel in left VTA/SN (a.u.: arbitrary units). Beta weights are estimated using a GLM including three boxcar function regressors associated with the three baseline reward conditions at the current block n, plus three stick function regressors associated with the different baseline reward conditions at the subsequent block $n + 1$. Beta weights are shown for displaying purposes, and no statistical tests were conducted on them. (D) Corresponding data plotted for left sGC. (E) Correlation between the individual behavioral parameter describing the effect of large (L) minus medium (M) and small (S) baseline reward at the current block on force and the neural activation for L minus M and S baseline reward in left VTA/SN (−8, −17, −13; $Z = 3.22$, $p = .011$ SVC). (F) Corresponding data plotted for left sGC (−8, 31, −15; $Z = 3.29$, $p = .024$ SVC).

Table 2. Neural Results Relative to the Main Analyses

Region	*Coordinates*	*t*	*p (uncorrected)*	*Z*	*p SVC*	*Cluster Size*	
Large Minus Medium and Small Baseline Reward at Current Block							
Right VTA/SN	9 −22 −15	2.72	.008	2.43	.095	0	
Left VTA/SN	−11, −22, −10	3.76	.001	3.07	.014*	20	
Right ventral striatum	17, 8, −5	3.25	.002	2.81	.042*	18	
Left ventral striatum	−13, 13, −3	2.30	.011	2.56	.095*	0	
Right sGC	7, 21, −10	−4.55	.001	−3.59	.005*	25	
Left sGC	−8, 31, −15	−6.20	<.001	−4.36	<.001*	22	
Right amygdala	17 6 −23	−2.93	.005	−2.58	.069	0	
Left amygdala	−21, 3, −25	−3.24	.003	−2.80	.044*	15	
Correlation with Pressing Force Exerted over Blocks							
Right VTA/SN	10 −12 −13	2.67	.008	2.41	.071	0	
Left VTA/SN	−8, −15, −15	2.98	.004	2.64	.032*	8	
Right ventral striatum	12, 6, −3	0.63	.236	0.72	.512	0	
Left ventral striatum	−8, 3, 0	0.51	.307	0.35	.563	0	
Right sGC	5, 28, −5	−3.29	.002	−2.86	.027*	21	
Left sGC	−1 31 −18	−3.03	.004	−2.50	.059	1	
Right amygdala	22 −2 −23	−2.42	.014	−2.21	.102	0	
Left amygdala	−21 −2 −20	−1.99	.031	−1.86	.177	0	

Region	*Coordinates*	*t*	*p (uncorrected)*	*Z*	*p SVC*	*Cluster Size*	*Pearson r*
Correlation across Subjects between the Behavioral Effect on Force and the Neural Contrast for Large Minus Medium and Small Baseline Reward at Current Block							
Right VTA/SN	10 −22 −18	2.91	.005	2.57	.071	0	0.590
Left VTA/SN	−8, −17, −13	3.90	.001	3.22	.011*	21	0.702
Right ventral striatum	12 6 −3	2.38	.015	2.17	.144	0	0.511
Left ventral striatum	−13 6 −3	3.31	.002	2.85	.068	12	0.637
Left sGC	−8, 31, −15	−4.01	.001	−3.29	.024*	25	−0.711
Right sGC	10, 26, −15	−3.03	.004	−2.65	.060	2	−0.603
Right amygdala	22, 1, −28	−1.98	.032	−1.86	.227	0	−0.404
Left amygdala	−21, 1, −23	−2.04	.029	−1.90	.213	0	−0.454
Correlation across Subjects between the Behavioral Effect on Force and the Interaction Parameter Extracted from the PPI Relative to Amygdala							
Right VTA/SN	10, −20, −10	1.54	.069	1.48	.263	0	0.359
Left VTA/SN	−13, −15, −13	1.03	.147	1.05	.393	0	0.249
Right ventral striatum	7, 18, 0	1.57	.066	1.51	.256	0	0.365
Left ventral striatum	−11, 8, −5	1.37	.090	1.34	.305	0	0.324
Left sGC	−3, 31, −15	−4.26	<.001	−3.43	.005*	26	−0.726

Table 2. (*continued*)

Region	*Coordinates*	*t*	*p (uncorrected)*	*Z*	*p SVC*	*Cluster Size*	
Right sGC	7, 31, −15	−1.52	.071	−1.47	.267	0	0.355
Right amygdala	15, −7, −20	−1.62	.0	−1.55	.242	0	0.375
Left amygdala	−16, −7, −20	−1.08	.138	−1.09	.380	0	0.260

For all ROIs, we report uncorrected t statistics and relative uncorrected p values, plus small volume corrected (SVC) Z statistics and corrected p values. Significant statistics have $p < .05$ SVC and are marked with asterisks. Cluster size is reported using a threshold of $p < .005$ uncorrected. Pearson correlation coefficients are also reported for correlation analyses.

Unlike hypotheses (a) and (b), hypothesis (c) predicts that a relationship between baseline reward and force is fully explained by the relationship between baseline reward and VTA/SN activity. To test this, we built a GLM relating the behavioral parameter (encoding the difference in force between large minus medium and small baseline reward) to an intercept parameter plus two regressors encoding differential brain activity for large minus medium and small baseline reward at the peak activation voxel in left VTA/SN and at the peak activation voxel in right sGC. Hypotheses (a) and (b), but not hypothesis (c), predict the intercept parameter would be different from zero, but results showed no such significant difference ($t(17) = -1.077$, $p = .298$). Conversely, parameters associated with activity in VTA/SN ($t(17) = 3.074$, $p = .008$) and right sGC ($t(17) = -2.150$, $p = .048$) were significantly different from zero. We also ran a similar analysis using a likelihood ratio test to compare this GLM with an equivalent GLM that lacked the intercept parameter and found a nonsignificant chi-square statistic ($\chi^2(1) = 0.61$, $p = .435$), indicating that adding the intercept parameter did not significantly improve the model fitting. Altogether these results provide support for a hypothesis that average reward influences activation in VTA/SN and sGC, and these in turn influence force.

On the basis of substantial evidence highlighting a central role for amygdala in coordinating Pavlovian behavior (Talmi et al., 2008; Fanselow & Gale, 2003; Davis, 1992), we hypothesized that this region would modulate the response of VTA/SN and sGC to baseline reward. To investigate this, we ran a PPI analysis taking as seed region the peak activation voxel in amygdala for the contrast large minus medium and small baseline reward in the current block (Figure 5B). We next built a second-level regression

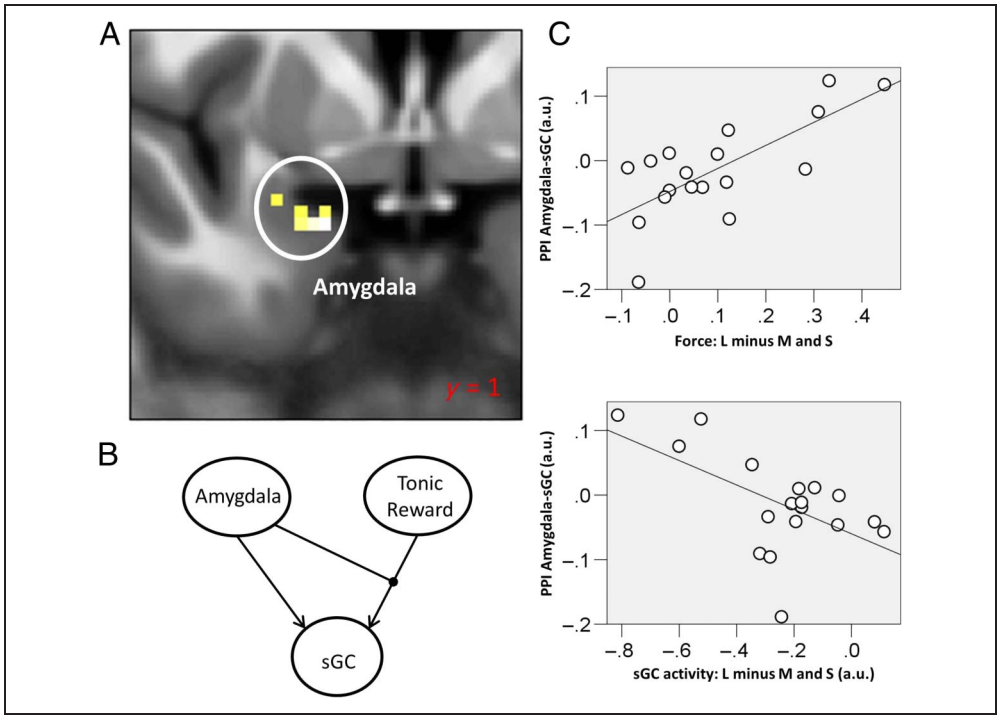

Figure 5. (A) Activation in left amygdala for the contrast small and medium minus large baseline reward of current block (−21, 3, −25; $Z = 2.80$, $p = .044$ SVC). (B) Schematic of a PPI analysis in which we tested whether the effect of baseline reward in the current block on sGC response was modulated by activity in the peak activation voxel in amygdala found when comparing large minus medium and small baseline reward in the current block. (C) In the top panel, the relationship between the behavioral parameter representing the strength of the behavioral effect on force (i.e., the effect of large (L) minus medium (M) and small (S) baseline reward at the current block) and the interaction parameter extracted from the PPI relative to amygdala and sGC (a.u.: arbitrary units) is plotted. Data are plotted for the peak activation voxel in sGC and show a significant correlation (−3, 31, −15; $Z = 3.43$, $p = .005$ SVC). In the bottom panel, the relationship between the parameter describing the effect of large minus medium and small baseline reward in the current block on sGC activity and the interaction parameter extracted from the PPI relative to amygdala and sGC is plotted. Data are plotted for the peak activation voxels in sGC extracted from the first and second analysis respectively and show a significant correlation ($r = -.573$; $p = .013$).

model of the PPI interaction parameter having as predictor the individual behavioral parameter representing the difference between large minus medium and small baseline reward at the current block. Across participants, the behavioral parameter correlated with the PPI interaction parameter in left sGC (Figure 5C; −3, 31, −15; $Z = 3.43$, $p = .005$ SVC) but not VTA/SN. Also, in left sGC, the PPI parameter inversely correlated with the parameter relative to the contrast large minus small and medium baseline reward at the current block (Figure 5C; $r = -.573$; $p = .013$). This suggests that, in participants showing a strong or weak behavioral effect, the amygdala was associated respectively with enhanced or attenuated sGC responses elicited with medium and small compared with large baseline reward at the current block.

DISCUSSION

Influential theories of motivation distinguish phasic and tonic components, the former linked to acquisition of new information about reward/punishment and the latter to repeated experiences of reward/punishment (Toates, 1986). To date, the neural substrates underlying phasic aspects of motivation have been characterized in terms of RPEs (D'Ardenne et al., 2008; Tobler et al., 2005; Schultz et al., 1997), and indeed, it is well established that areas such as VTA/SN and ventral striatum encode an RPE signal (Bartra et al., 2013; O'Doherty et al., 2003, 2004). However, previous empirical literature has not examined expression of a tonic motivation.

Here, based on theoretical (Niv et al., 2007) and prior empirical evidence (Beierholm et al., 2013; Guitart-Masip, Beierholm, et al., 2011), we set out to examine the relationship between brain activity and long-run reward rate, a concept connected with tonic aspects of motivation, in the absence of RPEs. Because in our paradigm the provision of new, surprising information (which leads to expression of an RPE) about average reward occurs well in advance of that average reward actually becoming available, we could decorrelate effects of an RPE from the effects of an average reward rate. In so doing, we show that average reward modulates a VTA/SN response which, in turn, was tightly coupled to the expression of a form of motivational vigor.

Although the BOLD signal is uninformative on the underlying neurochemical mechanisms, the likelihood that the effect we found in the VTA/SN involves the recruitment of dopamine circuitry is consistent with theories that propose a role for this neuromodulator in regulating motor vigor (Pessiglione et al., 2007; Salamone & Correa, 2002; Dickinson, Smith, & Mirenowicz, 2000; Berridge & Robinson, 1998). Physiological evidence indicates that tonic and phasic dopaminergic responses are at least partially independent, with the phasic signal linked to a bursting response and the tonic signal to the overall number of nonsilent dopaminergic neurons as well as presynaptic glutamatergic inputs (Lodge & Grace, 2006; Floresco, West, Ash, Moore, & Grace, 2003; Cheramy et al., 1990). Moreover, distinct regions projecting to VTA/SN affect these two forms of dopamine signaling in different ways. For example, inputs from the pedunculopontine nucleus influence primarily phasic bursts whereas inputs from ventral pallidum influence overall population activity (Lodge & Grace, 2006; Floresco et al., 2003; Cheramy et al., 1990). To identify both signals, we modeled the neural response to baseline reward with boxcar function regressors and the neural response to RPE with stick function regressors. We stress that it is not possible to infer the precise temporal profile of an underlying neural activation pattern with fMRI, and thus, whether the effect on BOLD activity found here is determined by tonic or phasic activity of neurons remains conjectural. However, previous research has coupled a phasic VTA/SN response almost exclusively with the expression of an RPE (D'Ardenne et al., 2008; Tobler et al., 2005; Schultz et al., 1997), and the lack of an RPE within blocks in our design renders it unlikely that a phasic, within-block, response is source of the effect on the block-wise BOLD signal we report here.

Evidence in favor of a somatotopic organization within VTA/SN is weak (Nambu, 2011), although a left/right differentiation depending on the nature of the target movement has been reported in dopaminergic regions (Gershman, Pesaran, & Daw, 2009). In addition, findings in rats and Parkinson's disease patients show an accentuated lateralized motor impairment linked with controlateral VTA/SN damage (Djaldetti, Ziv, & Melamed, 2006; Dunnett, Bjo, Stenevi, & Iversen, 1981). In our experiment, significant effects in VTA/SN were confined to the left hemisphere, whereas a mixed pattern was observed in other areas. However, we observed trends toward significance also on controlateral regions (see Table 2), and hence, a lack of significant effect might be explained by low power. This renders our data unsuitable to address any hypothesis about laterality in VTA/SN and ventral striatum.

An average reward rate could in principle be estimated either prospectively in a model-based manner or retrospectively in a model-free manner (Sutton & Barto, 1998). The latter estimate could be updated using a delta rule based on novel reward experience without reference to such considerations as the rules of the task. However, we found no evidence for such model-free learning as participants' behavior was unaffected by past experience of average reward. Instead, our data show that participants' behavior was influenced by the current baseline reward condition, suggesting that participants' performance reflected the sort of representation of task rules that model-based reinforcement learning algorithms exploit, albeit for average, rather than phasic, rewards.

Theoretical accounts relating average reward to vigor (Niv et al., 2007) suggest an instrumental effect, in that acting quickly is lucrative when there is a high prevailing

reward rate. However, one might hypothesize that evolution has favored increased vigor with increased reward rate even when this is not formally advantageous, as a Pavlovian effect. Our effect is of this nature because the baseline reward was delivered independently of performance and the performance-dependent reward was fixed. Such a Pavlovian effect might explain why reward (or its prediction) sometimes exerts an influence on behavior that might seem paradoxical (Rigoli, Pezzulo, & Dolan, 2016; Rigoli, Pavone, & Pezzulo, 2012; Dayan, Niv, Seymour, & Daw, 2006; Berridge & Robinson, 1998; Mackintosh, 1983; Williams & Williams, 1969). For instance, this is relevant in some forms of impulsivity in which performance decreases with reward in contexts where less vigorous behavior is more appropriate (Guitart-Masip, Fuentemilla, et al., 2011). By design, a larger baseline reward was associated with a decreased ratio between performance-contingent and noncontingent reward. In some contexts such as decision-making under risk, research has shown an endowment effect, which prescribes that agents treat the portion of reward dependent on choice/performance as larger when the portion of reward independent of performance is smaller (Kahneman et al., 1991). This would have led our participants to show increased vigor with smaller baseline reward, exactly contrary to our findings. An important difference between our task and other tasks in which an endowment effect has emerged is that, in the former, reward is delivered after an action is performed, whereas in the latter, the endowment is provided before making a choice. This might entail a different framing in such a way that a reward not yet collected might be attributed a higher weight, hence increasing Pavlovian vigor, whereas a reward collected already might increase a reward reference point and hence underweight the impact of future expected rewards.

Greater activation for large, compared with medium and small, average reward was seen in ventral striatum, linked to appetitive motivation. This region receives extensive dopaminergic inputs from VTA/SN and, as a consequence, is known to be robustly influenced by RPEs (Bartra et al., 2013; O'Doherty et al., 2003, 2004). Our data extend these previous observations by showing that activity in the two regions is coupled in response to average reward.

Enhanced sGC activity characterizes conditions linked to behavioral inhibition seen with negative mood states, such as depression and sadness (Drevets et al., 1997, 2008), whereas deep brain stimulation of this area (potentially having a deactivating effect) is reported as an effective treatment for refractory depression (Berlim et al., 2014; Mayberg et al., 2005). In this region, we observed a decreased activation for larger average reward. Moreover, sGC response to average reward influenced motor vigor consistent with a role for this area in behavioral inhibition. The sGC is known to be densely innervated with serotonergic inputs (Canli & Lesch, 2007). It has been suggested (Daw, Kakade, & Dayan, 2002; Deakin & Graeff, 1991) that some of the many effects of serotonin are opponent to dopamine, with dopamine being boosted by reward (and the attainment of safety; Oleson & Cheer, 2013) leading to an increase in vigor (Dayan, 2012) whereas (albeit based on substantially less evidence, and with contrary observations; Miyazaki et al., 2014) serotonin would be boosted by punishment (Schweimer & Ungless, 2010) and potential omission of reward leading to an increased inhibition (Cools, Roberts, & Robbins, 2008; Daw et al., 2002). There is no previous report of a phasic RPE in sGC, and asymmetries in the coding of reward and punishment (Boureau & Dayan, 2010) could be consistent with serotonin acting mainly in a tonic capacity, whereas dopamine would act in both phasic and tonic modes in its target regions.

The amygdala plays a key role in emotional regulation by coordinating Pavlovian responses and influencing Pavlovian instrumental transfer (Talmi et al., 2008; Corbit & Balleine, 2005; Fanselow & Gale, 2003). Consistent with an involvement of this structure in negative emotions (Sotres-Bayon, Sierra-Mercado, Pardilla-Delgado, & Quirk, 2012; Davis, 1992), we observed increased amygdala response with smaller average reward availability. A PPI analysis showed results consistent with the idea that, in participants showing a stronger Pavlovian vigor effect, the amygdala amplified the response to smaller and medium average reward in sGC while decreased such a response in participants showing a weaker Pavlovian vigor effect. However, we stress that PPI does not test for directionality, and therefore, these data are also consistent with the hypothesis that sGC exerts a modulatory influence on amygdala. These data extend to the appetitive domain previous animal reports consistent with the idea that amygdala coordinates Pavlovian fear responses by gating the expression of such responses that are more directly regulated by prefrontal regions (Sotres-Bayon et al., 2012).

In summary, we provide evidence that activity in VTA/SN increases with larger average reward when controlling for RPE and that this neural response enhances the expression of Pavlovian motor vigor. An opposite activation pattern was found in sGC, whose activation decreased with larger average reward and was associated with decreased motor vigor. The amygdala was found to amplify the sGC response to smaller average reward in participants with stronger behavioral vigor effect and to decrease such sGC response in participants with weaker behavioral vigor effect. Our findings shed important light on the neural substrates underlying tonic aspects of motivation.

Acknowledgments

This work was supported by the Wellcome Trust (Ray Dolan Senior Investigator Award 098362/Z/12/Z). The Wellcome Trust Centre for Neuroimaging was supported by core funding from the Wellcome Trust 091593/Z/10/Z. P. D. was supported by the

Gatsby Charitable Foundation. We are grateful for the advice on an earlier version of this paper from Tobias Hauser, Peter Smittenaar, and Zeb Kurth-Nelson.

Reprint requests should be sent to Francesco Rigoli, The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, 12 Queen Square, London, UK WC1N 3BG, or via e-mail: f.rigoli@ucl.ac.uk.

REFERENCES

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76,* 412–427.

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., et al. (2013). Dopamine modulates reward-related vigor. *Neuropsychopharmacology, 38,* 1495–1503.

Berlim, M. T., McGirr, A., Van den Eynde, F., Fleck, M. P., & Giacobbe, P. (2014). Effectiveness and acceptability of deep brain stimulation (DBS) of the subgenual cingulate cortex for treatment-resistant depression: A systematic review and exploratory meta-analysis. *Journal of Affective Disorders, 159,* 31–38.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28,* 309–369.

Boureau, Y. L., & Dayan, P. (2010). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology, 36,* 74–97.

Canli, T., & Lesch, K. P. (2007). Long story short: The serotonin transporter in emotion regulation and social cognition. *Nature Neuroscience, 10,* 1103–1109.

Cheramy, A., Barbeito, L., Godeheu, G., Desce, J. M., Pittaluga, A., Galli, T., et al. (1990). Respective contributions of neuronal activity and presynaptic mechanisms in the control of the in vivo release of dopamine. In M. B. H. Youdim & K. F. Tipton (Eds.), *Neurotransmitter actions and interactions* (pp. 183–193). Vienna: Springer.

Cools, R., Roberts, A. C., & Robbins, T. W. (2008). Serotoninergic regulation of emotional and behavioural control processes. *Trends in Cognitive Sciences, 12,* 31–40.

Corbit, L. H., & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *Journal of Neuroscience, 25,* 962–970.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science, 319,* 1264–1267.

Davey, C. G., Harrison, B. J., Yücel, M., & Allen, N. B. (2012). Regionally specific alterations in functional connectivity of the anterior cingulate cortex in major depressive disorder. *Psychological Medicine, 42,* 2071–2081.

Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience, 15,* 353–375.

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks, 15,* 603–616.

Dayan, P. (2012). Instrumental vigour in punishment and reward. *European Journal of Neuroscience, 35,* 1152–1168.

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks, 19,* 1153–1160.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313,* 684–687.

Deakin, J. W., & Graeff, F. G. (1991). 5-HT and mechanisms of defence. *Journal of Psychopharmacology, 54,* 305–315.

Dickinson, A., Smith, J., & Mirenowicz, J. (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behavioral Neuroscience, 114,* 468.

Djaldetti, R., Ziv, I., & Melamed, E. (2006). The mystery of motor asymmetry in Parkinson's disease. *The Lancet Neurology, 5,* 796–802.

Drevets, W. C., Price, J. L., Simpson, J. R., Todd, R. D., Reich, T., Vannier, M., et al. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature, 386,* 824–827.

Drevets, W. C., Savitz, J., & Trimble, M. (2008). The subgenual anterior cingulate cortex in mood disorders. *CNS Spectrums, 13,* 663.

Dunnett, S. B., Bjo, A., Stenevi, U., & Iversen, S. D. (1981). Behavioural recovery following transplantation of substantia nigra in rats subjected to 6-OHDA lesions of the nigrostriatal pathway. I. Unilateral lesions. *Brain Research, 215,* 147–161.

Fanselow, M. S., & Gale, G. D. (2003). The amygdala, fear, and memory. *Annals of the New York Academy of Sciences, 985,* 125–134.

Floresco, S. B., West, A. R., Ash, B., Moore, H., & Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience, 6,* 968–973.

George, M. S., Ketter, T. A., Parekh, P. I., Horwitz, B., Herscovitch, P., & Post, R. M. (1995). Brain activity during transient sadness and happiness in healthy women. *American Journal of Psychiatry, 152,* 341–351.

Gershman, S. J., Pesaran, B., & Daw, N. D. (2009). Human reinforcement learning subdivides structured action spaces by learning effector-specific values. *Journal of Neuroscience, 29,* 13524–13531.

Guitart-Masip, M., Beierholm, U. R., Dolan, R., Duzel, E., & Dayan, P. (2011). Vigor in the face of fluctuating rates of reward: An experimental examination. *Journal of Cognitive Neuroscience, 23,* 3933–3938.

Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J., Dayan, P., Dolan, R. J., et al. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *Journal of Neuroscience, 31,* 7867–7875.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives, 5,* 193–206.

Kohn, N., Falkenberg, I., Kellermann, T., Eickhoff, S. B., Gur, R. C., & Habel, U. (2014). Neural correlates of effective and ineffective mood induction. *Social Cognitive and Affective Neuroscience, 9,* 864–872.

Lodge, D. J., & Grace, A. A. (2006). The laterodorsal tegmentum is essential for burst firing of ventral tegmental area dopamine neurons. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 5167–5172.

Mackintosh, N. J. (1983). *Conditioning and associative learning* (p. 316). Oxford, UK: Clarendon Press.

Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., et al. (1999). Reciprocal limbic-cortical function and negative mood: Converging PET findings in depression and normal sadness. *American Journal of Psychiatry, 156,* 675–682.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., et al. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron, 45,* 651–660.

Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J., & Pessiglione, M. (2013). Neurocomputational account of how the human

brain decides when to have a break. *Proceedings of the National Academy of Sciences, U.S.A., 110,* 2641–2646.
Miyazaki, K. W., Miyazaki, K., Tanaka, K. F., Yamanaka, A., Takahashi, A., Tabuchi, S., et al. (2014). Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards. *Current Biology, 24,* 2033–2040.
Nambu, A. (2011). Somatotopic organization of the primate basal ganglia. *Frontiers in Neuroanatomy, 5,* 26.
Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology, 191,* 507–520.
O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304,* 452–454.
O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38,* 329–337.
Oleson, E. B., & Cheer, J. F. (2013). On the role of subsecond dopamine release in conditioned avoidance. *Frontiers in Neuroscience, 7,* 96.
Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., et al. (2007). How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science, 316,* 904–906.
Phan, K. L., Fitzgerald, D. A., Nathan, P. J., Moore, G. J., Uhde, T. W., & Tancer, M. E. (2005). Neural substrates for voluntary suppression of negative affect: A functional magnetic resonance imaging study. *Biological Psychiatry, 57,* 210–219.
Rauch, S. L., & Drevets, W. C. (2009). Neuroimaging and neuroanatomy of stress-induced and fear circuitry disorders. In G. Andrews, D. S. Charney, P. J. Siravotka, & D. A. Regie (Eds.), *Stress-induced and fear circuitry disorders: Refining the research agenda for DSM-V* (pp. 215–254). Arlington, VA: American Psychiatric Association.
Rigoli, F., Pavone, E. F., & Pezzulo, G. (2012). Aversive Pavlovian responses affect human instrumental motor performance. *Frontiers in Neuroscience, 6,* 134.
Rigoli, F., Pezzulo, G., & Dolan, R. J. (2016). Prospective and Pavlovian mechanisms in aversive behaviour. *Cognition, 146,* 415–425.
Salamone, J. D., & Correa, M. (2002). Motivational views of reinforcement: Implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research, 137,* 3–25.
Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599.
Schweimer, J. V., & Ungless, M. A. (2010). Phasic responses in dorsal raphe serotonin neurons to noxious stimuli. *Neuroscience, 171,* 1209–1215.
Skvortsova, V., Palminteri, S., & Pessiglione, M. (2014). Learning to minimize efforts versus maximizing rewards: Computational principles and neural correlates. *Journal of Neuroscience, 34,* 15621–15630.
Sotres-Bayon, F., Sierra-Mercado, D., Pardilla-Delgado, E., & Quirk, G. J. (2012). Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. *Neuron, 76,* 804–812.
Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. Cambridge, MA: MIT Press.
Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human Pavlovian–instrumental transfer. *Journal of Neuroscience, 28,* 360–368.
Toates, F. M. (1986). *Motivational systems (No. 4)*. CUP Archive. Cambridge: Cambridge University Press.
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science, 307,* 1642–1645.
Williams, D., & Williams, H. (1969). Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior, 12,* 511.