

Action controls dopaminergic enhancement of reward representations

Marc Guitart-Masip^{a,1}, Rumana Chowdhury^{a,b}, Tali Sharot^a, Peter Dayan^c, Emrah Duzel^{b,d}, and Raymond J. Dolan^a

^aWellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom; ^bInstitute of Cognitive Neuroscience, ^cGatsby Computational Neuroscience Unit, University College London, London, W1CN 4AR, United Kingdom; and ^dInstitute of Cognitive Neurology and Dementia Research, Otto von Guericke University Magdeburg, 39120 Magdeburg, Germany

Edited by Mortimer Mishkin, National Institute for Mental Health, Bethesda, MD, and approved March 26, 2012 (received for review February 7, 2012)

Dopamine is widely observed to signal anticipation of future rewards and thus thought to be a key contributor to affectively charged decision making. However, the experiments supporting this view have not dissociated rewards from the actions that lead to, or are occasioned by, them. Here, we manipulated dopamine pharmacologically and examined the effect on a task that explicitly dissociates action and reward value. We show that dopamine enhanced the neural representation of rewarding actions, without significantly affecting the representation of reward value as such. Thus, increasing dopamine levels with levodopa selectively boosted striatal and substantia nigra/ventral tegmental representations associated with actions leading to reward, but not with actions leading to the avoidance of punishment. These findings highlight a key role for dopamine in the generation of appetitively motivated actions.

functional MRI | striatum | decision-making

Substantial evidence indicates that the neuromodulator dopamine plays at least two roles in behavioral guidance. One is to signal prediction errors regarding expected reward value (1–3), prediction errors that are also used by the striatum to guide selection of appropriate actions (4–6). A second, less emphasized, role for dopamine is to invigorate actions associated with reward (7, 8). Thus, dopamine depletion results in decreased motor activity and decreased motivated behavior (9, 10) along with decreased vigor or motivation to work for rewards in demanding reinforcement schedules (7, 11). These joint roles in reward prediction and motivational control can create situations that result in conflict, such as what happens when reward is attainable solely by not acting (no-go).

It has been suggested that the exact opposite may be true of serotonin. That is, although the nature of any opponency between appetitive and aversive systems remains the subject of much debate (12–14), there are suggestions that serotonin acts as a mirror to dopamine and is associated with behavioral inhibition in aversive contexts (13–17). These joint roles would lead to conflict when punishment can be avoided only by acting (go).

Such conflicting cases have been a focal point of recent studies where the critical manipulation involved an orthogonalization of action requirements and outcome valence (15, 18). In the latter study we highlighted anticipatory representations in the striatum and substantia nigra/ventral tegmental area (SN/VTA) that reflected a dominance of action representation over an expected dominance of reward value representations, thereby hinting at a specific role for dopamine in motivation for action (18).

Here, we tested the contributions of dopamine and serotonin to action and valence representations, with a specific focus on the areas we had previously highlighted, namely the striatum and SN/VTA. First, we trained participants on a balanced 2 (reward/punishment) × 2 (go/no-go) task that orthogonalized action and valence (18). When performance reached 95% correct choices in all conditions, participants were then assigned to receive placebo, levodopa (150 mg), or citalopram (24 mg in oral drops, equivalent to 30 mg in tablets). The pharmacological agents are assumed to increase postsynaptic levels of dopamine and serotonin, respectively. We then tested the very same participants, under identical task conditions, while acquiring functional (f)MRI (18). A key difference between our protocol and those used in previous

studies addressing the relationship between action and valence (19, 20) is its ability to separate activity elicited by anticipation, action performance, and the receipt of an outcome. Thus, within this design we could index the effects of dopamine and serotonin manipulations on brain activity elicited during an anticipatory phase, separate from effects associated with action implementation (motor response) or outcome processing.

Our focus in this experiment is on the invigoration of an action before actual execution of any overt motor component. The process of action invigoration is likely to be multifaceted and includes deployment of cognitive resources (attention and sensory process) that allow a directing effect on the specific action that is being prepared. This dual association of motoric and cognitive components that interact to sculpt a motor response is a general mechanism that allows adaptive interactions with the environment. To what extent invigoration of action, and the associated deployment of distinct cognitive resources, can be specifically attributed to the observed anticipatory responses in the midbrain/basal ganglia network goes beyond the immediate goal and scope of the present study.

Results

Fifty-two subjects performed a go/no-go task involving reward and punishment (18) after administration of placebo (20), levodopa (16), or citalopram (16). On each trial, participants first saw one of four abstract fractal cues (Fig. 1 and *SI Materials and Methods*). These cues informed the subjects of a requirement to emit (go) or not emit a button press (no-go) to a target within a time limit of 700 ms as well as the valence of feedback related to an appropriately executed response to the target (reward/no reward or punishment/no punishment). The target was a circle that could appear either to the right or to the left, and participants had to indicate its location with a manual button press. Thus, there were four trial types signaled by the identity of a fractal cue presented at the beginning of the trial comprising (i) press the correct button to the target to gain a reward (go to win), (ii) press the correct button to the target to avoid punishment (go to avoid losing), (iii) do not press any button to the target to gain a reward (no-go to win), and (iv) do not press any button to the target to avoid punishment (no-go to avoid losing). In half the trials, the fractal image was not followed by the target and no feedback was provided (Fig. 1). Therefore, at the beginning of each trial, fractal images specified action requirements (go vs. no-go) and outcome valence (reward vs. punishment). However, the actual target detection task and outcome receipt were presented only in half the trials. In effect this manipulation allowed us to decorrelate activity related to anticipation cued by fractal image presentation

Author contributions: M.G.-M., P.D., E.D., and R.J.D. designed research; M.G.-M. and R.C. performed research; M.G.-M. and T.S. analyzed data; and M.G.-M., T.S., P.D., E.D., and R.J.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: m.guitart@ucl.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1202229109/-DCSupplemental.

pation of action (go/no-go) or to anticipation of valence (win/avoid losing) we examined blood oxygen level-dependent (BOLD) activity evoked by fractal cue onset. By design these cues predict, on each trial, precise action requirements and their valence dependencies. Hence, unless stated otherwise all our analyses pertain to cue-related brain responses.

We first assessed main effects (action and valence) across all treatment groups to highlight voxels that responded to these experimental factors. We then assessed the impact of our pharmacological manipulations on these voxels. Note we failed to observe any behavioral effect with administration of citalopram and, in the interests of parsimony, we focus on levodopa treatment effects (for detailed results of the citalopram condition see Figs. S5 and S6).

A voxel-based analysis (Fig. S2) revealed a simple main effect of “action anticipation” (go > no-go) in a cluster that included peaks in right [Montreal Neurological Institute (MNI) space coordinates (x, y, z) 10, 12, 8; peak Z = 6.29; $P < 0.001$ family wise error corrected (FWE)] and left caudate (MNI -8, 7, 4; peak Z = 5.68; $P = 0.001$ FWE) as well as in left (MNI -28, -5, -3; peak Z = 5.65; $P = 0.001$ FWE; MNI -24, -1, 8; peak Z = 5.6; $P = 0.001$ FWE) and right putamen (MNI 26, 7, -4; peak Z = 5.36; $P = 0.004$ FWE; MNI 22, 6, 8; peak Z = 5.02; $P = 0.02$ FWE). This cluster also included the left thalamus (MNI -12, -20, 10; peak Z = 6.07; $P < 0.001$ FWE) extending into the bilateral SN/VTA [MNI 12, -18, -10; peak Z = 4.89; $P < 0.001$, small volume correction (SVC); MNI -7, -22, -14; peak Z = 4.3; $P = 0.002$, SVC]. We also observed an effect in bilateral insula (MNI 32, 26, 5; peak Z = 5.19; $P = 0.009$ FWE; MNI -42, 15, 0; peak Z = 4.89; $P = 0.034$ FWE) and bilateral cerebellum, respectively (MNI 38, -46, -32; peak Z = 5.37; $P = 0.004$ FWE; MNI -6, -74, -18; peak Z = 4.87; $P = 0.037$ FWE). This overall pattern is strikingly similar to what we have reported previously (18).

Consistent with previous studies (18, 22, 23) we also found greater activity in the winning relative to avoid losing conditions in right ventral striatum (MNI 8, 18, -2; peak Z = 4.56; $P = 0.001$ SVC FWE) and right SN/VTA (MNI 7, -20, 15; peak Z =

3.64; $P = 0.018$ SVC FWE) (Fig. S2 C and D). However, the effect of “valence” (win > avoid losing) was both weaker and more restricted than the main effect of action, and closer inspection shows it was largely driven by a difference between win and avoid-losing trials in go but not in no-go trials (Fig. S3).

We next extracted parameter estimates in voxels showing a simple main effect of action across all treatment groups, for each participant in each condition (go to win, go to avoid losing, no-go to win, and no-go to avoid losing). These voxels were then constrained using anatomical region of interest (ROI) masks (bilateral caudate, putamen, SN/VTA, insula, and cerebellum), such that the mean values of each parameter estimate across the activated voxels in each anatomical structure could be calculated separately. In the striatum and the SN/VTA, levodopa enhanced the difference in BOLD response relative to placebo between go and no-go conditions, an effect evident solely when the possible trial outcome was a win (Fig. 3). This effect was significant in an action \times valence \times treatment interaction in all functional ROIs within the striatum and the SN/VTA. The sole exception was right caudate and left SN/VTA, which showed an action \times treatment and an action \times valence interaction (Table 1). In the insula and the left cerebellum there was no interaction (Table 1).

Thus, the effects of levodopa on action representation were restricted to striatum and SN/VTA. Post hoc *t* tests showed that within the striatum levodopa resulted in more positive parameter estimates, relative to placebo, for the go to win conditions and more negative parameter estimates for the no-go to win condition ($P < 0.05$). This result contrasts with the pattern seen in SN/VTA, where a difference was entirely driven by an increase in parameter estimates in go to win conditions, with no effect on the no-go to win condition ($P < 0.01$).

We next ascertained parameter values for each condition, averaged across all voxels for the regions showing a main effect of valence. In the right ventral striatum, this analysis highlighted greater activity in win relative to avoid losing trials but only when a go action was required to obtain an outcome [action \times valence

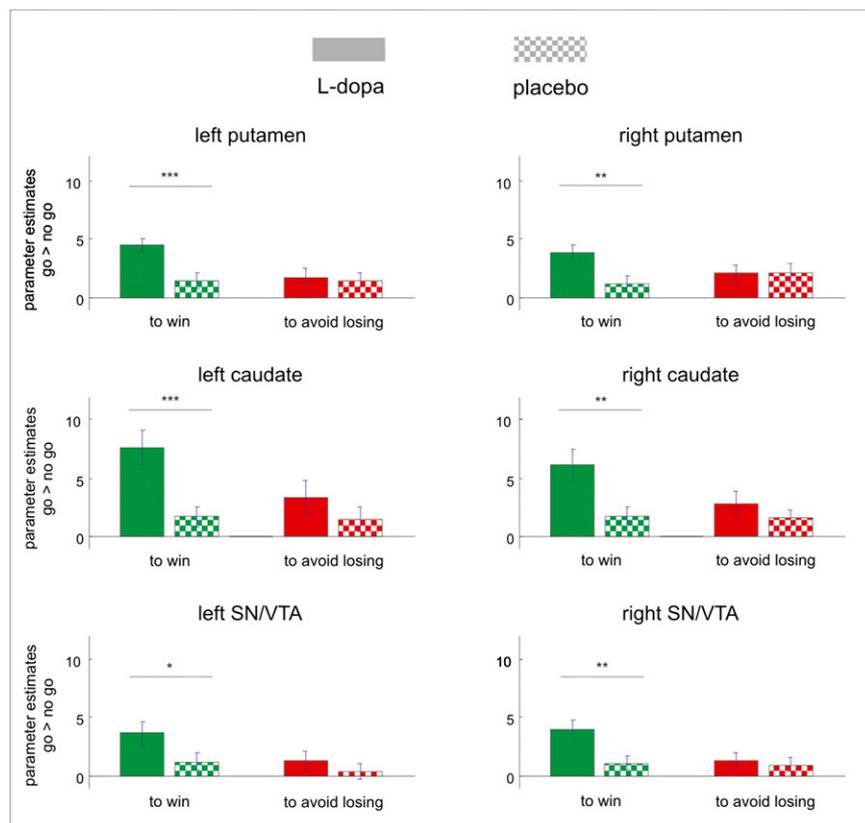


Fig. 3. Effects of levodopa on action anticipation in the striatum and the SN/VTA. Mean (\pm SEM) parameter estimates for the contrast between fractal images indicating go trials and fractal images indicating no-go trials (action anticipation) are presented separately for win (green) and avoid losing (red) trials. Solid color represents the effects for the group that received levodopa and checkered color shows those for the group that received placebo. Regions of interest were functionally defined as activated voxels in the contrast “go > no-go” across all treatment groups within the caudate, the putamen, and the SN/VTA bilaterally. In all functional ROIs, Levodopa can be seen to increase activity in the contrast between go and no-go trials when the outcome of the correct action choice was a win. The contrast between go and no-go trials when the outcome of the correct action choice was avoidance of a loss was not affected by the drug. Post hoc comparisons were implemented by means of a *t* test: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.005$.

Table 1. Effects of valence and drug on action representation (go > no go) within voxels responsive to action

	Action by drug	Action by valence	Action by valence by drug
Left putamen	F(1,34) = 4.42; P = 0.043*	F(1,34) = 6.73; P = 0.014*	F(1,34) = 6.85; P = 0.014*
Right putamen	F(1,34) = 2.12; P = 0.154	F(1,34) = 0.74; P = 0.395	F(1,34) = 6.28; P = 0.017*
Left caudate	F(1,34) = 6.96; P = 0.012*	F(1,34) = 6.93; P = 0.013*	F(1,34) = 5.57; P = 0.024*
Right caudate	F(1,34) = 6.78; P = 0.014*	F(1,34) = 4.93; P = 0.033*	F(1,34) = 3.79; P = 0.06*
Left SN/VTA	F(1,34) = 4.14; P = 0.05*	F(1,34) = 4.84; P = 0.035*	F(1,34) = 0.96; P = 0.335
Right SN/VTA	F(1,34) = 4.5; P = 0.041*	F(1,34) = 7.21; P = 0.011*	F(1,34) = 6.29; P = 0.017*
Left insula	F(1,34) = 1.96; P = 0.661	F(1,34) = 0.07; P = 0.796	F(1,34) = 0.62; P = 0.436
Right insula	F(1,34) = 0.01; P = 0.941	F(1,34) = 3.65; P = 0.064	F(1,34) = 1.57; P = 0.221
Left cerebellum	F(1,34) = 0.97; P = 0.332	F(1,34) = 1.24; P = 0.274	F(1,34) = 2.32; P = 0.137
Right cerebellum	F(1,34) = 7.31; P = 0.011*	F(1,34) = 2.99; P = 0.093	F(1,34) = 2.17; P = 0.150

*Significant effects at $P < 0.05$.

interaction $F(1, 49) = 7.45$; $P = 0.010$]. This effect was not influenced by administration of levodopa [valence \times treatment $F(2, 49) = 0.29$, NS; and action \times treatment \times valence interaction $F(2, 49) = 1.30$, NS] (Fig. 4), suggesting it is not dopamine dependent. In the right SN/VTA, greater activity for win trials relative to lose trials was observed only in the go condition and only in participants that received levodopa [action \times valence \times treatment interaction $F(2, 49) = 15.37$; $P < 0.001$] (Fig. 4). Note that this cluster overlapped with the bigger cluster showing a main effect of action and an interaction between action, valence, and drug (compare Fig. S2 B and D), where levodopa specifically increased activity in the go to win condition.

Note that our design enabled us to disambiguate anticipatory BOLD elicited by the fractal images from the actual motor performance required to the targets. This disambiguation was possible as on 50% of trials the fractal images were not followed by the response targets. By including a regressor that accounted for the motor responses to the target, we can explain away variance associated with motor execution. To ensure the effects of levodopa on BOLD signals are not confounded by a drug-induced speeding of motor responses, we repeated the same analysis but then included reaction time of the go responses as a parametric regressor of no interest. The inclusion of this further regressor explains away variance due to the trial-by-trial fluctuations of motor performance revealed in reaction times. This analysis revealed a valence-by-drug-by-action interaction in the striatum, including the left putamen [$F(1, 34) = 6.02$; $P = 0.019$], the right putamen [$F(1, 34) = 8.77$; $P = 0.006$], and the left caudate [$F(1, 34) = 5.63$; $P = 0.023$], as well as in the right SN/VTA [$F(1, 34) = 8.74$; $P = 0.006$]. For details, see Tables S2 and S3 and Fig. S4. These results show that the reported effects do not have a trivial motor explanation. However, they do not imply that anticipatory brain responses are not related to performance of the motor responses. Indeed one likely possibility that we cannot test in the current dataset is that if all trials had included a target, then reaction times to this target might have been significantly correlated with regional BOLD responses to the fractals on a trial-by-trial basis.

For the sake of completeness, we note that these conclusions are identical when we include the citalopram group in the analysis (Figs. S5 and S6). Finally, activity to outcomes (rather than anticipation in response to cues) showed no modulation by either drug, as reported in Fig. S7.

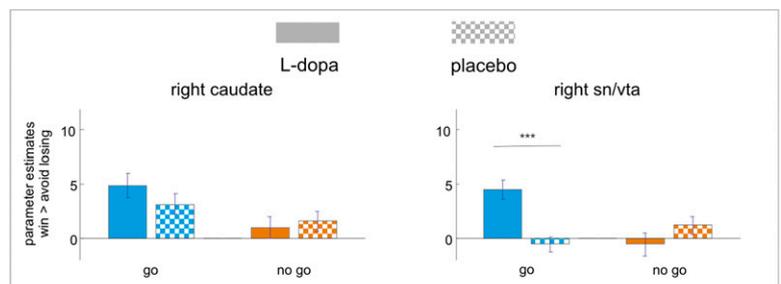
Discussion

We showed that a pharmacologically induced increase in brain dopamine levels resulted in faster behavioral responses, without impacting on the ability to retrieve and perform accurately go/no-go choices that subjects learned before drug administration. Neurally, this behavioral speeding was mirrored by increased activity for anticipated go vs. no-go choices in the striatum and the SN/VTA. Critically, these effects were observed only when the outcome of a potential action entailed a reward.

The treatment effects of levodopa show that dopamine can enhance neural representations of rewarded actions independent of any effect on the representation of reward value. The most straightforward effect of levodopa was observed in the go to win condition, a condition that combines an anticipation of action and an expectation of reward. Indeed, this conjunction of action and valence is the standard association in studies that have addressed the neural substrates of reward processing (e.g., refs. 4, 23, and 24). This conjunction also extends to Pavlovian paradigms, where rewards are delivered no matter what the subject actually does, but where hard-wired action patterns (such as anticipatory licking) are usually elicited by reward expectation (25–27). As expected, levodopa led to a speeding up in reaction times and also significantly increased the BOLD signal in dorsal and ventral striatum and the SN/VTA.

Although one would not expect, a priori, that increased availability of dopamine in the terminal field of dopamine neurons would impact on BOLD response in the SN/VTA, studies using genetically dopamine-deficient mice show that endogenously released dopamine plays a critical role in afferent control of dopamine neuron bursting (28). Given that the BOLD response is likely to reflect, at least in part, strength of presynaptic input to a given region (29), it is possible that administration of levodopa

Fig. 4. Effects of levodopa on valence anticipation in the striatum and the SN/VTA. Mean (\pm SEM) parameter estimates for the contrast between fractal images indicating win trials and fractal images indicating avoidance of loss trials (valence anticipation) are presented separately for go (blue) and no-go (orange) trials. Solid color represents the effects for the group that received levodopa and checkered color those for the group that received placebo. Regions of interest were functionally defined as activated voxels in the contrast “win > avoid losing” across all treatment groups within the caudate, the putamen, and the SN/VTA bilaterally. No significant differences between levodopa and placebo groups were found in the right caudate (including the nucleus accumbens), where an effect of valence anticipation was observed only in go trials. In the right SN/VTA, levodopa increased the contrast between win and avoid losing trials only when an action was required. Post hoc comparisons were implemented by means of a t test: *** $P < 0.005$.



results in increased afferent drive into the neurons located in the SN/VTA. Confirming this hypothesis, and identifying the neurons receiving such input, remains a task for future research.

In our experiment the case of no-go to win is especially interesting, given a previously reported dominance of action representation in the striatum (18). Here boosting dopamine actually decreased the overall striatal BOLD signal, the exact opposite of what might naively be expected from the valence association (win) of this condition. It would be of particular interest to explore the differential activity of direct and indirect pathways in the striatum, because the association of phasic dopamine with anticipation of reward poses particular problems for theories that couple dopamine increases and decreases directly with influence on go and no-go pathways, respectively (30). By contrast with its effects in the striatum, levodopa did not modulate BOLD responses for no-go to win cues in the SN/VTA, raising an intriguing question about the precise nature of a coding for expected gain.

In the go to avoid losing condition, there was no detectable BOLD correlate for the significant speeding in reaction times. One possibility is that this behavioral effect arises from a tonic increase in dopamine (as suggested in ref. 11), which would have been invisible to event-related fMRI. Levodopa is known to exert dual effects on dopamine release (31), with an increased availability of presynaptic dopamine increasing both phasic and tonic dopamine (32). These two signals are associated with different roles: tonic dopamine with motivational vigor associated with instrumental actions (7, 8, 11) and phasic dopamine with expression of a learning-based temporal difference prediction error for future reward (1, 2). It is not immediately clear how increases in either form of dopamine release would be expressed in the striatum or the SN/VTA BOLD signal (for a discussion, see ref. 33), so any conclusions must be tentative.

However, the question of involvement of phasic dopamine in the go to avoid losing condition is important. Two-factor theories of active avoidance behavior (34–36) suggest that the transition from an unsafe to a safe state is coded similarly to a reward, an idea central to various notions about vigor and learning in avoidance (13). The lack of an effect of levodopa in the go to avoid losing condition is troubling for this account. However, our results do not imply that dopamine is not involved in a go to avoid losing condition. In showing that dopamine involvement in the go to win condition is a dominant effect, our data does not outrule the possibility that future experiments, involving variations of our experimental paradigm and/or other experimental modalities including experiments with animals, might detect a dopaminergic involvement in the anticipation of actions leading to avoidance of a loss. It is known that fewer dopamine neurons respond to cues associated with punishment or punishment avoidance than respond to reward, and despite the fact that such units have long been recognized (37) they have only recently attracted significant attention (27, 38). Further, because many dopamine neurons are inhibited by punishments (39), and by an expectation of punishment, the net signal in the entire population may be subthreshold. Experimental manipulations involving single-unit recording that orthogonalize action and valence in a manner similar to the present experiment would be extremely important in addressing these mechanistic questions.

Beyond the effects of levodopa, we replicate previous findings (18) that when a requirement for action and valence outcome are orthogonalized, action representation dominates valence representation in dorsal striatum and SN/VTA (in the placebo condition). As we found before (18), there was no evidence of a modulation by the possibility of winning vs. avoiding losing in these regions. This result again bolsters the idea that the striatum signals a propensity to perform a given action independently from the value of the states in which that action is taken, akin to computational accounts of the purest form of actor in an actor-critic architecture (40).

However, the large number of subjects in the present study enabled us to see a robust main effect of anticipation of valence based on the cues (i.e., go to win + no-go to win > go to avoid losing + no-go to avoid losing) in ventral striatum, a signal that was only

marginally significant in the previous study. This signal is expected from a reward prediction error account (4–6, 41), although the lack of any task in half the anticipatory trials would render a signal that distinguishes predictions from prediction errors hard to discern. Importantly, this main effect was observed only when an action was required. This pattern of response in the ventral striatum suggests it reports action-dependent reward prediction error signals, i.e., expected reward conditional upon an action. The same was true for the SN/VTA, although this response depended critically on the effects of levodopa. In fact, a requirement for action associated with this signal is consistent with an extensive literature reporting reward prediction errors for state values under experimental conditions that control for action requirements indirectly through the use of explicit foil actions (see for example refs. 5, 6, and 42). In those studies, reward prediction errors were isolated by comparing actions leading to rewards with foil actions that did not result in reward.

Our results suggest limits to the generality of salience theories of dopamine (43), notwithstanding the recent imaging studies where evoked striatal activity did not differentiate between reward and punishment (19, 44, 45) or data that highlight punishment-associated dopamine neurons (27, 38, 46). As for pure prediction error accounts, saliency theories founder on our finding that responses to reward and punishment predictive cues are markedly different in these same brain areas, dependent on whether or not an active response is required, in keeping with the idea that anticipatory striatal representations are dominated by action rather than action-independent salience representations.

We did not find any behavioral effect of citalopram in our task. Serotonin has been associated with both punishment (12) and behavioral inhibition (13, 14, 17). One previous study found that tryptophan depletion abolishes punishment-induced inhibition akin to the disadvantage of performing a go response in the avoid losing condition compared with the winning condition we observe in our task (15). However, involvement of serotonin in inhibition is typically complicated (14, 47), and even the regional effects on serotonin concentration of single doses of citalopram are controversial (48).

From a methodological perspective the findings we report highlight the importance of a simultaneous manipulation of expected reward along with instrumental requirements to fully reveal the contribution of dopamine (or indeed any other neuromodulator). By implementing such an approach we highlight a dopaminergic contribution to the control of motivation in instrumental responding. Our results suggest that dopamine has an intimate involvement in signaling and invigorating actions likely to lead to a reward. Consequently, our findings help bind computational ideas regarding the role of dopamine that highlight a role of motivational vigor with emerging evidence for a dominance of action representations in the striatum.

Materials and Methods

Subjects. Fifty-two healthy volunteers were recruited through University College London (UCL). Participants were randomly assigned to one treatment group: 16 participants received levodopa (6 females, age range 18–34 y; mean 23.3, SD = 5 y), 16 participants received citalopram (11 females, age range 19–35 y; mean 23.6, SD = 4.67 y), and 20 participants received placebo (6 females, age range 19–27 y; mean 23, SD = 2.47 y). The study was double blind. All participants were right-handed and had normal or corrected-to-normal visual acuity. None of the participants reported a history of neurological, psychiatric, or any other current medical problems. All experiments were run with each subject's written informed consent and according to the local ethics clearance (University College London).

Experimental Design and Task. Each trial consisted of three events: a fractal cue, a target detection task, and an outcome. The target detection task involved spatial discrimination as the target was a circle displayed on one side of the screen for 1500 ms. When the target appeared, participants had the opportunity to press a button within a time limit of 700 ms to indicate the target side for go trials or not to press for no-go trials. The requirement to make a go or a no-go response was dependent on the preceding fractal cue. The trial timeline is displayed in Fig. 1 (*SI Materials and Methods*).

Participants completed the task inside the scanner 1 h after receiving levodopa (150 mg + 37.5 mg benserazide) or 3 h after receiving citalopram (24 mg in drops, which is equivalent to 30 mg in tablet). Participants completed

a subjective state analog-scales questionnaire on three occasions. We did not detect any difference in subjective ratings between experimental conditions (*SI Materials and Methods*, Table S4).

Scanning was divided into four 8-min sessions comprising 20 trials per condition (*SI Materials and Methods*). Before any treatment was administered, we ensured that subjects learned the meaning of the fractal images and performed the task with a high level of accuracy (*SI Materials and Methods*). Subjects were paid £40 for participation. Moreover, they were told that they would be paid their earnings from the task up to a maximum of £20.

Behavioral Data Analysis. The behavioral data were analyzed using the statistics software SPSS, version 16.0. The number of correct on-time button press responses per condition was analyzed with a $2 \times 2 \times 2$ ANOVA with action (go/no-go) and valence (win/lose) as within-subjects factors and treatment (levodopa/placebo) as a between-subject factor. Response speed in “go” trials was analyzed by considering the button press RTs to targets and the proportion of trials where button press RTs exceeded the response deadline. To further analyze these effects we performed post hoc *t* tests.

fMRI Data Acquisition and Analysis. fMRI was performed on a 3-Tesla Siemens Allegra scanner with echo planar imaging (EPI) with BOLD contrast. Standard

preprocessing was performed and for each individual subject we estimated a general linear model that included our four conditions of interest (*SI Materials and Methods*).

At first-level analysis, regionally specific condition effects were tested by using linear contrasts for each subject and each condition of interest. We tested the effects of the pharmacological treatment on voxels that were sensitive to the contrast of interest across pharmacological groups. We defined functional ROIs, at the second level, by using a 2×2 ANOVA with the factors “action” (go/no-go) and valence (win/lose) that did not include treatment as a factor. Results of this selection process are reported FWE corrected for the whole brain or for small volume in areas of interest at $P < 0.05$ (*SI Materials and Methods*).

Within functional ROIs, mean parameter estimates for each condition and subject were extracted across activated clusters, using the marsbar toolbox (49), and analyzed using an ANOVA with action, valence, and drug (*SI Materials and Methods*).

ACKNOWLEDGMENTS. We thank Claire Burley for assistance in data acquisition. This work was supported by a Wellcome Trust Programme Grant 078865/Z/05/Z (to R.J.D.), the Gatsby Charitable Foundation (P.D.), a Marie Curie Fellowship PIEF-GA-2008-220139 (to M.G.-M.), and the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 779, TP A7).

- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- O’Doherty J, et al. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454.
- O’Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- Salamone JD, Correa M, Mingote SM, Weber SM (2005) Beyond the reward hypothesis: Alternative functions of nucleus accumbens dopamine. *Curr Opin Pharmacol* 5: 34–41.
- Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28:309–369.
- Ungerstedt U (1971) Adipsia and aphagia after 6-hydroxydopamine induced degeneration of the nigro-striatal dopamine system. *Acta Physiol Scand Suppl* 367: 95–122.
- Palmiter RD (2008) Dopamine signaling in the dorsal striatum is essential for motivated behaviors: Lessons from dopamine-deficient mice. *Ann N Y Acad Sci* 1129: 35–46.
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.
- Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616.
- Boureau YL, Dayan P (2011) Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36:74–97.
- Cools R, Nakamura K, Daw ND (2011) Serotonin and dopamine: Unifying affective, motivational, and decision functions. *Neuropsychopharmacology* 36:98–113.
- Crockett MJ, Clark L, Robbins TW (2009) Reconciling the role of serotonin in behavioral inhibition and aversion: Acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J Neurosci* 29:11993–11999.
- Dayan P, Huys QJ (2009) Serotonin in affective control. *Annu Rev Neurosci* 32:95–126.
- Soubrie P (1986) Reconciling the role of central serotonin neurons in human and animal behavior. *Behav Brain Sci* 9:319–364.
- Guitart-Masip M, et al. (2011) Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J Neurosci* 31:7867–7875.
- Tricomi EM, Delgado MR, Fiez JA (2004) Modulation of caudate activity by action contingency. *Neuron* 41:281–292.
- Elliott R, Newman JL, Longe OA, William Deakin JF (2004) Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *Neuroimage* 21:984–990.
- Gray JA, McNaughton M (2000) *The Neuropsychology of Anxiety: An Inquiry into the Function of the Septohippocampal System* (Oxford Univ Press, Oxford), 2nd Ed.
- Haber SN, Knutson B (2009) The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* 35(1):4–26.
- Knutson B, Adams CM, Fong GW, Hommer D (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159.
- Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol* 84:3072–3077.
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902.
- Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645.
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- Paladini CA, Robinson S, Morikawa H, Williams JT, Palmiter RD (2003) Dopamine controls the firing pattern of dopamine neurons via a network feedback mechanism. *Proc Natl Acad Sci USA* 100:2866–2871.
- Logothetis NK (2008) What we can do and what we cannot do with fMRI. *Nature* 453: 869–878.
- Frank MJ, Seeberger LC, O’Reilly RC (2004) By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Grace AA (1991) Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience* 41:1–24.
- Breitenstein C, et al. (2006) Tonic dopaminergic stimulation impairs associative learning in healthy subjects. *Neuropsychopharmacology* 31:2552–2564.
- Düzel E, et al. (2009) Functional imaging of the human dopaminergic midbrain. *Trends Neurosci* 32:321–328.
- Mowrer OH (1947) On the dual nature of learning: A reinterpretation of conditioning and problem solving. *Harv Educ Rev* 17:102–148.
- Maia TV (2010) Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn Behav* 38:50–67.
- Moutoussis M, Bental RP, Williams J, Dayan P (2008) A temporal difference account of avoidance learning. *Network* 19:137–160.
- Mirenowicz J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–451.
- Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci USA* 106: 4894–4899.
- Ungless MA, Magill PJ, Bolam JP (2004) Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303:2040–2042.
- Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442: 1042–1045.
- D’Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008) BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319:1264–1267.
- Redgrave P, Prescott TJ, Gurney K (1999) Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 22:146–151.
- Carter RM, MacInnes JJ, Huettel SA, Adcock RA (2009) Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Front Behav Neurosci*, 10.3389/fnro.08.021.2009.
- Wrase J, et al. (2007) Different neural systems adjust motor behavior in response to reward and punishment. *Neuroimage* 36:1253–1262.
- Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron* 68:815–834.
- Druke B, et al. (2010) Serotonergic modulation of response inhibition and re-engagement? Results of a study in healthy human volunteers. *Hum Psychopharmacol* 25:472–480.
- Bari A, et al. (2010) Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* 35:1290–1301.
- Brett M, Anton J-L, Valabregue R, Poline J-B (2002) Region of interest analysis using an SPM toolbox. *8th International Conference on Functional Mapping of the Human Brain*. Available on CD-ROM in *NeuroImage* 16:2.