

# Action Dominates Valence in Anticipatory Representations in the Human Striatum and Dopaminergic Midbrain

Marc Guitart-Masip,<sup>1,2</sup> Lluís Fuentemilla,<sup>1</sup> Dominik R. Bach,<sup>2</sup> Quentin J. M. Huys,<sup>2,3</sup> Peter Dayan,<sup>3</sup> Raymond J. Dolan,<sup>2</sup> and Emrah Duzel<sup>1,4</sup>

<sup>1</sup>Institute of Cognitive Neuroscience, University College London, London, W1CN 4AR, United Kingdom, <sup>2</sup>Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom, <sup>3</sup>Gatsby Computational Neuroscience Unit, University College London, London, W1CN 4AR, United Kingdom, and <sup>4</sup>Institute of Cognitive Neurology and Dementia Research, Otto von Guericke University, 39120 Magdeburg, Germany

The acquisition of reward and the avoidance of punishment could logically be contingent on either emitting or withholding particular actions. However, the separate pathways in the striatum for go and no-go appear to violate this independence, instead coupling affect and effect. Respect for this interdependence has biased many studies of reward and punishment, so potential action–outcome valence interactions during anticipatory phases remain unexplored. In a functional magnetic resonance imaging study with healthy human volunteers, we manipulated subjects' requirement to emit or withhold an action independent from subsequent receipt of reward or avoidance of punishment. During anticipation, in the striatum and a lateral region within the substantia nigra/ventral tegmental area (SN/VTA), action representations dominated over valence representations. Moreover, we did not observe any representation associated with different state values through accumulation of outcomes, challenging a conventional and dominant association between these areas and state value representations. In contrast, a more medial sector of the SN/VTA responded preferentially to valence, with opposite signs depending on whether action was anticipated to be emitted or withheld. This dominant influence of action requires an enriched notion of opponency between reward and punishment.

## Introduction

In instrumental conditioning, particular outcomes are realized, or obviated, through discrete action choices, controlled by outcome valence. Rewarded (appetitive) action choices are repeated and punished (aversive) action choices are deprecated, although the nature of the opponency between appetitive and aversive systems remains the subject of debate (Gray and McNaughton, 2000). Aside from valence or affect opponency between reward and punishment, a key role in instrumental conditioning is also played by a logically orthogonal spectrum of effect, spanning invigoration to inhibition of action (Gray and McNaughton, 2000; Niv et al., 2007; Boureau and Dayan, 2011; Cools et al., 2011). This effect spectrum is enshrined in the structure of parts of the striatum that are involved in instrumental control, in which partially segregated direct and indirect pathways are de-

scribed for go (invigoration) and no-go (inhibition), respectively (Gerfen, 1992; Frank et al., 2004).

Although instrumental behavior thus seems to arise through an interaction of valence and action spectra, our understanding of their association remains partial. There is evidence for a close coupling of reward and go and some evidence for a coupling between punishment and no-go (Gray and McNaughton, 2000). In contrast, there is intense theoretical debate concerning how instrumental behavior is generated for the opposite associations, namely reward–no-go and punishment–go (Gray and McNaughton, 2000).

A conventional coupling between reward and go responses in human functional neuroimaging studies on instrumental conditioning has led to important findings, such as an encoding of various forms of temporal difference prediction errors for future reinforcement in the ventral and dorsal striatum (O'Doherty et al., 2004) and the identification of brain regions engaged in anticipation of wins and losses (Delgado et al., 2000; Knutson et al., 2001; Guitart-Masip et al., 2010). Overall, these studies have contributed to a view that the striatum, especially its ventral subdivision, and the midbrain regions harboring dopamine neurons are associated with the representation of rewards, prediction errors for rewards, and reward-associated stimuli (Haber and Knutson, 2010). However, in these experiments, the requirement to act (i.e., to go) is typically constant, and so a possible organizational principle of the striatum along an action spectrum has not been fully explored. Thus, in this study, we examined valence together with anticipation of a requirement to either act or in-

Received Dec. 7, 2010; revised April 4, 2011; accepted April 5, 2011.

Author contributions: M.G.-M., L.F., D.R.B., Q.J.M.H., P.D., R.J.D., and E.D. designed research; M.G.-M. and L.F. performed research; M.G.-M. analyzed data; M.G.-M., P.D., R.J.D., and E.D. wrote the paper.

The authors declare no competing financial interests.

This work was supported by Wellcome Trust Programme Grant 078865/Z/05/Z (R.J.D.), the Gatsby Charitable Foundation (P.D.), Marie Curie Fellowship PIEF-GA-2008-220139 (M.G.-M.), and Deutsche Forschungsgemeinschaft SFB 779 and TP A7. We thank Dr. Chris Lambert, Dr. Jörn Diedrichsen, and Dr. John Ashburner for discussion about midbrain normalization, Dr. Chloe Hutton for help with physiological noise correction, and Dr. Tali Sharot, Dr. Estela Camera, Dr. Molly Crockett, and Dr. Regina Lopez for comments on previous versions of this manuscript.

Correspondence should be addressed to Marc Guitart-Masip, Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, WC1N 3AR, UK. E-mail: m.guitart@ucl.ac.uk.

DOI:10.1523/JNEUROSCI.6376-10.2011

Copyright © 2011 the authors 0270-6474/11/317867-09\$15.00/0

hibit action, thereby disentangling these from an associated appetitive or aversive outcome delivery.

We orthogonalized action and valence in a balanced 2 (reward/punishment) × 2 (go/no-go) design. A key difference between our protocol and those of previous studies addressing the relationship between action and valence (Elliott et al., 2004; Tricomi et al., 2004) is that it allowed us to separate activity elicited by anticipation, action performance, and obtaining an outcome. Thus, unlike previous experiments, we could analyze outcome valence and action effects during anticipation as separate factors. We focused our analysis on the striatum and the putatively dopaminergic midbrain because of the close association between this neuro-modulator, reward, go, and indeed vigor (Schultz et al., 1997; Berridge and Robinson, 1998; Salamone et al., 2005; Niv et al., 2007).

## Materials and Methods

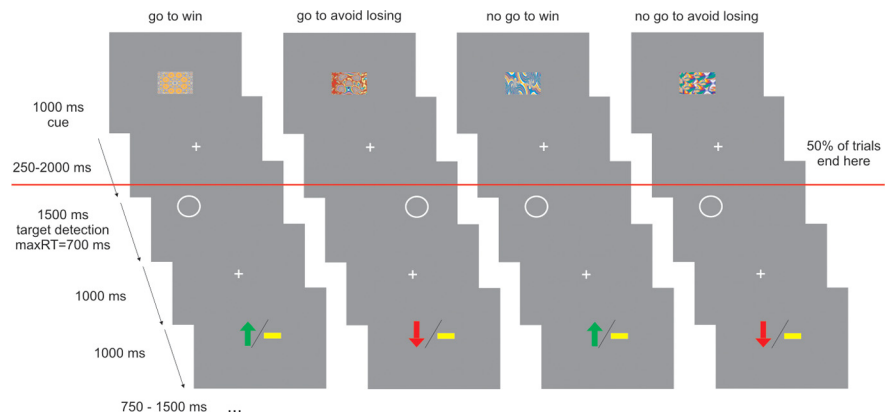
**Subjects.** Eighteen adults participated in the experiment (nine female and nine male; age range, 21–27 years; mean ± SD, 23 ± 1.72 years). All participants were healthy, right-handed, and had normal or corrected-to-normal visual acuity. None of the participants reported a history of neurological, 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, London, UK).

**Experimental design and task.** The goal of our experimental design was to disentangle neural activity related to the anticipation of action and valence. To investigate the relationship between the two predominant spectra in ventral and dorsal striatum, we had to include both punishment (losses) and reward (gains), along with go and no-go. With one notable exception (Crockett et al., 2009), the bulk of the human literature into instrumental conditioning has focused on rewards that are only available given an overt response (O'Doherty, 2004; Daw et al., 2006). These are well aligned with a tight coupling between reward and invigoration and thus do not address our critical questions. Alternatively, studies including punishment have systematically included a motor response as a means of its avoidance (Delgado et al., 2000; Knutson et al., 2001) but did not include no-go conditions.

Similarly, in other studies that address the role of action or salience on reward processing, subjects had to perform a motor response as part of the task (Zink et al., 2003, 2004; Elliott et al., 2004; Tricomi et al., 2004). Although explicit foil actions were used to control for the overall requirement to act, they did not study the case of controlled no-go, for which a lack of action itself constitutes the instrumental requirement.

Finally, it is important to note that it is not possible merely to use a comparison between classical and instrumental conditioning. Although in classical conditioning experiments, rewards or punishments are obtained without regard to a motor response, this form of conditioning is associated with the generation of conditioned anticipatory responses such as licking, approach, salivation, etc. These anticipatory responses, which generally result in increased biological efficiency in the interaction between an organism and unconditioned stimuli (Domjan, 2005), can in principle confound any attempt to isolate pure anticipation of valence.

Our trials consisted of three events: a fractal cue, a target detection task, and an outcome. The trial timeline is displayed in Figure 1. In each trial, subjects saw one of four abstract fractal cues for 1000 ms. The fractal cues indicated, first, whether the participant would subsequently be required to emit a button press (go) or omit a button press (no-go), in the target detection task. The cues also indicated the



**Figure 1.** Experimental design. On each trial, one of four possible fractal images indicated the combination between action (making a button press in go trials or withholding a button press in no-go trials) and valence at outcome (win or lose). Actions were required in response to a circle that followed the fractal image after a variable delay. In go trials, subjects indicated via a button press on which side of the screen the circle appeared. In no-go trials, they withheld a response. After a brief delay, outcome was signaled in which a green upward arrow indicated a win of £1, a downward red arrow indicated a loss of £1, and a horizontal bar indicated the absence of a win or a loss. In go to win trials, a correct button press was rewarded. In go to avoid losing trials, a correct button press avoided punishment. In no-go to win trials, withholding a button press led to reward. In no-go to avoid losing trials, withholding a button press avoided punishment. The outcome was probabilistic so that 70% of correct responses were rewarded in win trials, and 70% of correct responses were not punished in lose trials. The red line indicates that half of the trials did not include the target detection task and the outcome. Subjects were trained in the task and fully learned the contingencies between the different fractal images and task requirements before scanning.

potential valence of the outcome related to performance in the target detection task (reward/no reward or punishment/no punishment). After a variable interval (250–2000 ms) after offset of the fractal image, the target detection task started. The target was a circle displayed on one side of the screen for 1500 ms. At this point, 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. At 1000 ms after the offset of the circle, subjects were presented with the outcome implied by their response. The outcome was presented for 1000 ms: a green arrow pointing upward meant they had won £1, a red arrow pointing downwards meant that they had lost £1, and a yellow horizontal bar indicated they did not win or lose any money. The outcome was probabilistic so that 70% of correct responses were rewarded in win trials, and 70% of correct responses were not punished in lose trials.

Thus, there were four trial types depending on the nature of the fractal cue presented at the beginning of the trial: (1) press the correct button in the target detection task to gain a reward (“go to win”); (2) press the correct button in the target detection task to avoid punishment (“go to avoid losing”); (3) do not press a button in the target detection task to gain a reward (“no-go to win”); and (4) do not press a button in the target detection task to avoid punishment (“no-go to avoid losing”).

Critically, on half the trials, target detection and outcome were omitted (Fig. 1). Therefore, at the beginning of the trial, fractal images specified action requirements (go vs no-go) and outcome valence (reward vs punishment), but the actual target detection and potential delivery of an outcome only happened in half the trials. We implemented this manipulation because it allowed us to decorrelate activity related to an anticipation phase cued by the fractal stimuli from activity related to actual motor performance in the target detection task and obtaining an outcome. One additional benefit of this design is that we could avoid the suboptimality of having to introduce long jitters between distinct task components. If every trial had been followed by the target detection task, anticipation of action would have been followed by action execution and anticipation of inaction by action inhibition in all correct trials. This would have resulted in highly correlated regressors for the anticipation and execution or withholding of a motor response, making it impossible

to separate activity elicited by anticipation, action performance, and the delivery of an outcome.

Scanning was divided into four 8 min sessions comprising 20 trials per condition, 10 trials in which the target detection task and the outcome was displayed and 10 trials in which only the fractal image was displayed. Subjects were told that they would be paid their earnings from the task up to a maximum of £35. To ensure that subjects learned the meaning of the fractal images and performed the task correctly during the scanning, we instructed them as to the meaning of each fractal image before the actual scanning began. Moreover, subjects performed one block of the task with 10 trials per condition in which the outcome of each trial also included text providing feedback whether the executed response was correct or not and whether the response was on time. Finally, after this initial training session and before actual scanning, subjects performed another run of the task that was identical to the task performed during the scanning. This ensured that subjects experienced the possibility of the absence of the target detection task. Therefore, the presence of trials without target detection and outcome was not surprising during the crucial acquisition of functional magnetic resonance imaging (fMRI) data. Both training sessions were performed inside the scanner while the structural scans were acquired.

**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 two-way repeated-measures ANOVA with action (go/no-go) and valence (win/lose) as factors. Response speed in go trials was analyzed by considering the button press reaction times (RTs) to targets and the proportion of trials in which button press RTs exceeded the response deadline. To further analyze these effects, we performed *post hoc* *t* tests.

**fMRI data acquisition.** fMRI was performed on a 3 tesla Siemens Allegra magnetic resonance scanner with echo planar imaging (EPI). Functional data were acquired in four scanning sessions containing 117 volumes with 40 slices, covering a partial volume that included the striatum and the midbrain (matrix, 128 × 128; 40 oblique axial slices per volume angled at −30° in the anteroposterior axis; spatial resolution, 1.5 × 1.5 × 1.5 mm; TR, 4000 ms; TE, 30 ms). The fMRI acquisition protocol was optimized to reduce susceptibility-induced blood oxygen level-dependent (BOLD) response sensitivity losses in inferior frontal and temporal lobe regions (Weiskopf et al., 2006). Six additional volumes at the beginning of each series were acquired to allow for steady-state magnetization and were subsequently discarded. Anatomical images of each subject's brain were collected using multi-echo 3D fast, low-angle shot sequence (FLASH) for mapping proton density, T1 and magnetization transfer (MT) at 1 mm<sup>3</sup> resolution, and by T1-weighted inversion recovery prepared EPI sequences (spatial resolution, 1 × 1 × 1 mm). Additionally, individual field maps were recorded using a double-echo FLASH sequence (matrix size, 64 × 64; 64 slices; spatial resolution, 3 × 3 × 3 mm; gap, 1 mm; short TE, 10 ms; long TE, 12.46 ms; TR, 1020 ms) for distortion correction of the acquired EPI images. Using the FieldMap toolbox, field maps were estimated from the phase difference between the images acquired at the short and long TE.

**fMRI data analysis.** Data were analyzed using SPM8 (Wellcome Trust Centre for Neuroimaging, University College London). Preprocessing included realignment, unwrapping using individual field maps, and spatial normalization to the Montreal Neurological Institute (MNI) space with spatial resolution after normalization of 1 × 1 × 1 mm. We used the unified segmentation algorithm available in SPM to perform normalization. This has been shown to achieve good intersubject coregistration for brain areas such as caudate, putamen, and brainstem (Klein et al., 2009). Moreover, successful coregistration of the substantia nigra/ventral tegmental area (SN/VTA) was also checked by manually drawing a region of interest (ROI) for each subject, in native space, and inspecting the overlap of ROIs after applying the same normalization algorithm (data not shown). Finally, data were smoothed with a 6 mm FWHM Gaussian kernel. The fMRI time series data were high-pass filtered (cutoff, 128 s) and whitened using an AR(1)-model. For each subject, a statistical model was computed by applying a canonical hemodynamic response function combined with time and dispersion derivatives.

Our 2 × 2 factorial design included four conditions of interest that were modeled as separate regressors in a general linear model (GLM): go

to win trials, go to avoid losing trials, no-go to win trials, and no-go to avoid losing trials. We also modeled the onset of the target detection task separately for trials in which subjects performed a button press and for trials in which subjects did not perform a button press, respectively; the onset of the outcome, which could be win £1, lose £1, or no monetary consequences. Finally, we modeled separately the onsets of fractal images that were followed by incorrect performance. Note that the model used to analyze the data pooled together neutral outcomes from win trials (go to win and no-go to win conditions) together with neutral outcomes from lose trials (go to avoid losing and no-go to avoid losing conditions). Because the values of outcomes are assessed relative to expectations and the neutral outcomes have different effects if the alternative outcome is a win or a loss, the resulting analysis cannot be optimal for characterizing brain responses to the outcomes. This is because the goal of the present work was to study brain responses during the anticipatory phase and the experimental design, together with the GLM, optimized the detection of brain responses to the fractal images. To capture residual movement-related artifacts, six covariates were included (the three rigid-body translation and three rotations resulting from realignment) as regressors of no interest. Regionally specific condition effects were tested by using linear contrasts for each subject and each condition (first-level analysis). The resulting contrast images were entered into a second-level random-effects analysis. For the anticipatory phase, the hemodynamic effects of each condition were assessed using a 2 × 2 ANOVA with the factors "action" (go/no-go) and valence (win/lose). For the outcome onset, we assessed the hemodynamic effect of each condition using a one-way ANOVA with valence as a factor (win, lose, or neutral).

Results are reported familywise error (FWE) corrected for small volume in areas of interest at  $p < 0.05$ . The predicted activations in the midbrain and the striatum were tested using small volume correction (SVC) using anatomically defined regions of interest: the striatum as whole, the ventral striatum, and the SN/VTA of the midbrain (main origin of dopaminergic projections). The striatum as a whole ROI was defined using Marsbar (Brett et al., 2002) and included the caudate and the putamen. The ventral striatum ROI was drawn with Marsbar as two spheres of 8 mm around the coordinates referred to as right [MNI space coordinates (shown as *x,y,z* throughout), 11.11, 11.43, −1.72] and left (MNI space coordinates, −11.11, 11.43, −1.72) nucleus accumbens in previous publication (Knutson et al., 2005). This resulted in an ROI that incorporated the nucleus accumbens and ventral striatum as described in a recent review (Haber and Knutson, 2010). The SN/VTA ROI was manually defined, using the software MRICro and the mean MT image for the group. On MT images, the SN/VTA can be distinguished from surrounding structures as a bright stripe (Bunzeck and Düzél, 2006). It should be noted that, in primates, reward-responsive dopaminergic neurons are distributed across the SN/VTA complex, and it is therefore appropriate to consider the activation of the entire SN/VTA complex rather than, a priori, focusing on its subcompartments such as the VTA (Düzél et al., 2009). For this purpose, a resolution of 1.5 mm<sup>3</sup>, as used in the present experiment, allows sampling over 200 voxels of the SN/VTA complex, which has a volume of 350–400 mm<sup>3</sup>. This does not imply that the whole complex responds as a unit, and we have previously highlighted (Düzél et al., 2009) the possible existence of gradients in the functional anatomy of the SN/VTA in nonhuman primates (Haber et al., 2000) and the usefulness of high-resolution imaging of the entire SN/VTA to detect these functional gradients (Düzél et al., 2009).

## Results

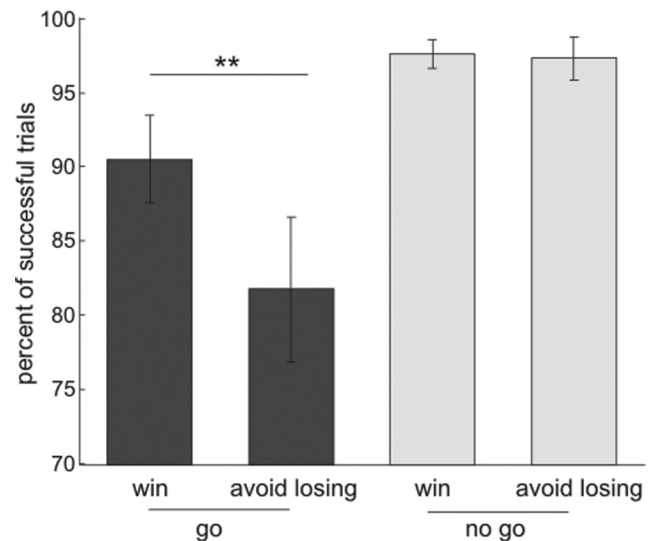
### Anticipation of losses impairs task performance when action is required

A two-way repeated-measures ANOVA on the percentage of successful target response trials, with action (go/no-go) and valence (win/lose) as factors, revealed a main effect of action ( $F_{(1,17)} = 22.88$ ,  $p < 0.001$ ), a main effect of valence ( $F_{(1,17)} = 13.2$ ,  $p = 0.002$ ), and an action × valence interaction ( $F_{(1,17)} = 12.28$ ,  $p = 0.003$ ). As illustrated in Figure 2A, anticipation of punishment decreased the percentage of successful (correct on time response to targets)

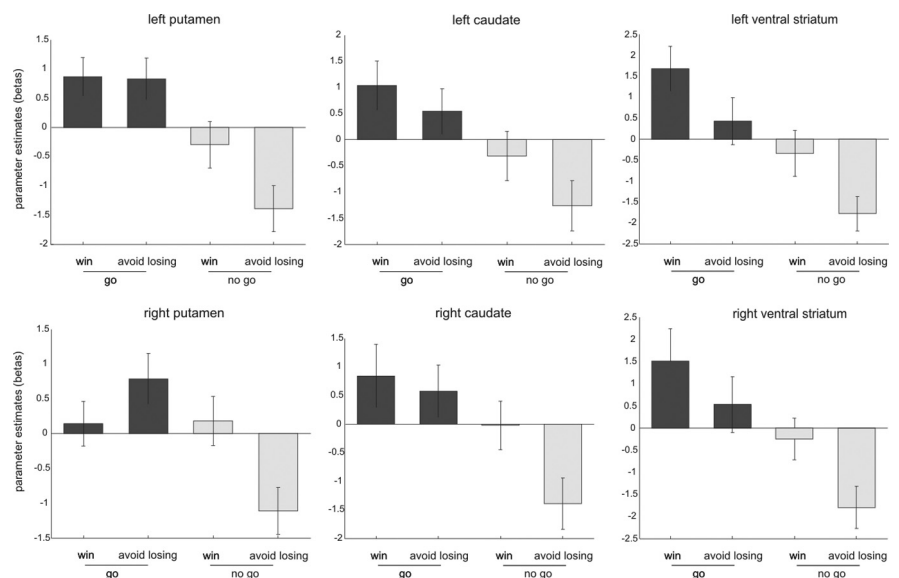
trials in the go conditions (repeated-measures Student's *t* test,  $t_{(17)} = 3.79$ ,  $p = 0.001$ ) but did not affect task performance in no-go conditions ( $t_{(17)} = 0.33$ , NS). Note that errors in the go trials included incorrect no-go responses and RTs that exceeded the requisite response window (700 ms). The percentage of incorrect (no-go) responses in go trials was higher for the lose condition (mean  $\pm$  SEM percentage of incorrect no-go responses for the win condition,  $1.11 \pm 0.36$ ; for the lose condition,  $4.17 \pm 1.43$ ;  $t_{(17)} = 2.5$ ,  $p = 0.023$ ). The percentage of trials in which RTs exceeded the response deadline in go trials was also higher for the lose condition (mean  $\pm$  SEM percentage of trials with too slow responses for the go to win trials,  $8.06 \pm 1.75$ ; for the go to avoid losing trials,  $13.89 \pm 2.69$ ;  $t_{(17)} = 2.61$ ,  $p = 0.018$ ). Furthermore, mean RTs were slower for correct go responses in the lose condition (mean  $\pm$  SEM RT for go to win trials,  $529.24 \pm 13.5$ ; mean  $\pm$  SEM RT for go to avoid losing trials  $557.81 \pm 18.1$ ;  $t_{(17)} = 3$ ,  $p = 0.008$ ). Thus, despite high levels of response accuracy throughout the scanning session (correct responses  $>95\%$  for all conditions), anticipation of loss had a negative impact on task performance whenever a go response was required. There was no evidence for a similar effect of valence in the no-go condition, whereas anticipation of gains exerted no deleterious effect on an ability to withhold responses in no-go trials. These data are strongly indicative of a behavioral asymmetry between actions for gains and losses.

### Anticipatory brain responses for action and valence

We focused our fMRI analysis on responses evoked by the onset of fractal images because these cues predicted both valence (win/lose) and response requirement (go/no-go) in each trial. To examine whether the striatum responded to action anticipation, valence, or both, we conducted an ROI analysis on this region using a second-level two-way ANOVA with action (go/no-go) and valence (win/lose) as factors within anatomically defined ROIs in the striatum. All six ROIs within the striatum (for details, see Fig. 3, Table 1) showed a main effect of action but no effect of valence. Only in the right putamen did we find an action  $\times$  valence interaction, an effect driven by action effects (a difference between go and no-go) in the lose conditions but none in the win conditions. To increase the power of our analysis, we pooled the data from all striatal ROIs and performed a three-way ANOVA with ROI (six different subdivisions), action (go/no-go), and valence (win/lose). This revealed a main effect of action alone ( $F_{(1,17)} = 11.87$ ,  $p = 0.001$ ) without any main effect of valence ( $F_{(1,17)} = 2.21$ ,  $p = 0.155$ ) or any action  $\times$  valence interaction ( $F_{(1,17)} = 1.18$ ,  $p = 0.292$ ). These results demonstrate in an unbiased manner that, in our paradigm, action anticipation was widely represented within the striatum. This contrasted with the absence of significant valence anticipation effects. Although the second part of this general conclusion is based on a failure to reject the null hypothesis, it is nevertheless important to highlight the contrast with the consistent difference between the go to win and the



**Figure 2.** Behavioral results. Mean percentage of trials in which subjects did a correct response within the response deadline for the go trials (blue) and did not emit any response on the no-go trials (red). *Post hoc* comparisons were implemented by means of repeated-measures *t* test. \*\* $p < 0.005$ .



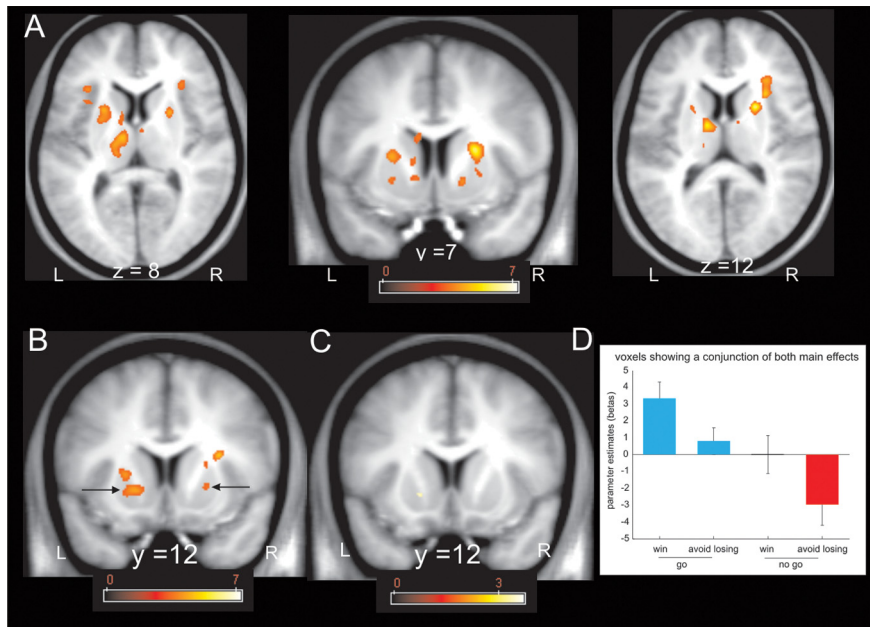
**Figure 3.** Response to anticipation of action and valence within anatomically defined ROIs in the striatum. Fractal images indicating go trials elicited higher activity than fractal images indicating no-go trials in all three bilateral ROIs: caudate, putamen, and ventral striatum (main effect of action,  $p < 0.05$ ) (for details, see Table 1). In the right putamen, fractal images indicating go trials elicited higher activity than fractal images indicating no-go trials only in the lose conditions (action  $\times$  valence interaction,  $p < 0.05$ ) (for details, see Table 1).

no-go to win conditions. These two conditions had the same value expectation, but *post hoc* pairwise *t* tests showed that they elicited markedly different BOLD responses in left putamen ( $t_{(17)} = 2.22$ ,  $p = 0.04$ ) and left ventral striatum ( $t_{(17)} = 2.69$ ,  $p = 0.016$ ). In the left caudate and right ventral striatum, this difference between the go to win and the no-go to win conditions approached significance ( $t_{(17)} = 1.9$ ,  $p = 0.075$  and  $t_{(17)} = 1.8$ ,  $p = 0.089$ , respectively). Conversely, we emphasize that none of the pairwise comparisons between the go to win and the go to avoid losing conditions was significant.

We next conducted a whole-brain, voxel-based analysis that revealed a simple main effect of action (go  $>$  no-go) in three local

**Table 1. Summary results within anatomically defined ROIs in the striatum**

	Main effect of action	Main effect of valence	Action × valence interaction
Right caudate	$F_{(1,17)} = 7.06; p = 0.017$	$F_{(1,17)} = 2.02; p = 0.17$	$F_{(1,17)} = 1.09; p = 0.31$
Left caudate	$F_{(1,17)} = 11.74; p = 0.003$	$F_{(1,17)} = 1.35; p = 0.26$	$F_{(1,17)} = 0.3; p = 0.66$
Right putamen	$F_{(1,17)} = 5.78; p = 0.03$	$F_{(1,17)} = 0.64; p = 0.44$	$F_{(1,17)} = 5.9; p = 0.027$
Left putamen	$F_{(1,17)} = 16.4; p = 0.001$	$F_{(1,17)} = 1.25; p = 0.28$	$F_{(1,17)} = 2.41; p = 0.14$
Right ventral striatum	$F_{(1,17)} = 9.37; p = 0.007$	$F_{(1,17)} = 2.83; p = 0.11$	$F_{(1,17)} = 2.16; p = 0.65$
Left ventral striatum	$F_{(1,17)} = 16.29; p = 0.001$	$F_{(1,17)} = 2.97; p = 0.1$	$F_{(1,17)} = 0.048; p = 0.83$



**Figure 4.** Voxel-based results within the striatum in response to anticipation of action and valence. **A**, Fractal images indicating go trials elicited higher dorsal striatal activity than fractal images indicating no-go trials ( $p < 0.001$  uncorrected;  $p < 0.05$  SVC within the whole-striatum ROI). The color scale indicates  $t$  values. **B**, Fractal images indicating go trials elicited higher ventral putamen activity than fractal images indicating no-go trials ( $p < 0.001$  uncorrected;  $p < 0.05$  SVC within the ROI restricted to the ventral striatum). The color scale indicates  $t$  values. **C**, Fractal images indicating that win trials elicited higher ventral putamen activity than fractal images indicating lose trials ( $p < 0.001$  uncorrected; did not survive SVC within the ROI restricted to the ventral striatum). The color scale indicates  $t$  values. **D**, Parameter estimates at the peak coordinates confirm that activation of the ventral putamen signals the anticipation of action. Although the anticipation of valence also seems to have an effect in the left ventral putamen, the effect did not survive SVC. Coordinates are given in MNI space. Error bars indicate SEM (note that these parameter estimates were not used for statistical inference). L, Left; R, right.

maxima within dorsal striatum that survived SVC within the anatomical whole-striatum ROI (Fig. 4A). These foci were located in the right putamen (MNI space coordinates, 23, 7, 12; peak  $Z$  score, 4.92;  $p = 0.001$  FWE), right caudate (MNI space coordinates, 21, 7, 13; peak  $Z$  score, 4.75;  $p = 0.003$  FWE), and left putamen (MNI space coordinates, -23, 11, 13; peak  $Z$  score, 4.07;  $p = 0.04$  FWE). The first two belonged to a single cluster that extended between the right caudate and putamen but was segregated into a caudate and putamen portion in the ROI analysis because the dividing internal capsule white matter tract, which separates these structures, was not part of the ROI. When we constrained our analysis to an ROI restricted to the ventral striatum (Fig. 4B), we found significant action anticipation-related activation in the left (MNI space coordinates, -17, 12, -2; peak  $Z$  score, 3.99;  $p = 0.007$  FWE) and right (MNI space coordinates, 16, 7, -5; peak  $Z$  score, 3.71;  $p = 0.018$  FWE) ventral putamen.

In keeping with previous studies of reward (Delgado et al., 2000; Knutson et al., 2001; O'Doherty et al., 2002), the only striatal region showing a main effect of valence (win > lose) was located in the left ventral putamen (MNI space coordinates, -17,

12, -39) (Fig. 4C). However, this main effect only approached significance when the search volume was restricted to the ventral striatum (peak  $Z$  score, 3.23;  $p = 0.076$  FWE). Because this cluster overlapped with the cluster showing a main effect of action, we extracted betas for the conjunction cluster (Fig. 4D). Even in this ventral striatal cluster, the dominant activity pattern was an effect of action (go > no-go), with greater activity in go to win compared with the no-go to win condition. The difference between the go to win and go to avoid losing on one hand, and the no-go to win and the no-go to avoid losing on the other hand, is reminiscent of the previously reported valence effects in the fMRI literature in which action requirements were not manipulated (Delgado et al., 2000; Knutson et al., 2001). A weak effect of valence, however, is compatible with recent evidence that ventral striatal activation to wins and losses is less differentiable than individual valence responses compared with neutral trials (Wrase et al., 2007; Cooper and Knutson, 2008). Note that all our experimental conditions were highly salient by virtue of their affective significance, and, on this basis, we do not consider the signals we find are likely to reflect mere salience (Redgrave et al., 1999). An intriguing possibility is that increased ventral striatum activity in the go to win relative to go to

avoid losing condition might be related to our behavioral finding of better performance in the go to win compared with the go to avoid losing condition.

Midbrain activity (Fig. 5A, B) showed a simple main effect of action (go > no-go) within a left lateral region of SN/VTA that survived SVC within our a priori ROI (MNI space coordinates, -12, -19, -7; peak  $Z$  score, 3.33;  $p = 0.039$  FWE). This contrasted with the response profile within a right medial SN/VTA region (Fig. 5C, D), which showed a significant interaction of action and valence that survived SVC within our a priori ROI (MNI space coordinates, 8, -9, -10; peak  $Z$  score, 3.85;  $p = 0.008$  FWE), with anticipation of action inducing activation in win trials but deactivation in lose trials. These findings also survived physiological noise correction for cardiac and respiratory phases (data not shown). This dissociable pattern is strikingly similar to findings from a recent electrophysiological study in monkeys (Matsumoto and Hikosaka, 2009), which distinguished between the response profiles of two distinct groups of dopaminergic neurons. One group, located in dorsolateral substantia nigra/VTA complex, re-

sponded to both reward and punishment predictive cues, whereas the other, located more ventromedially, responded preferentially to reward-predictive stimuli. We note that heterogeneity within dopaminergic midbrain is also described at a cellular level (Lammel et al., 2008) and in rat electrophysiological recordings (Brischoux et al., 2009), although the anatomical location of dopamine neurons responsive to punishment within the SN/VTA complex might differ between rats and monkeys.

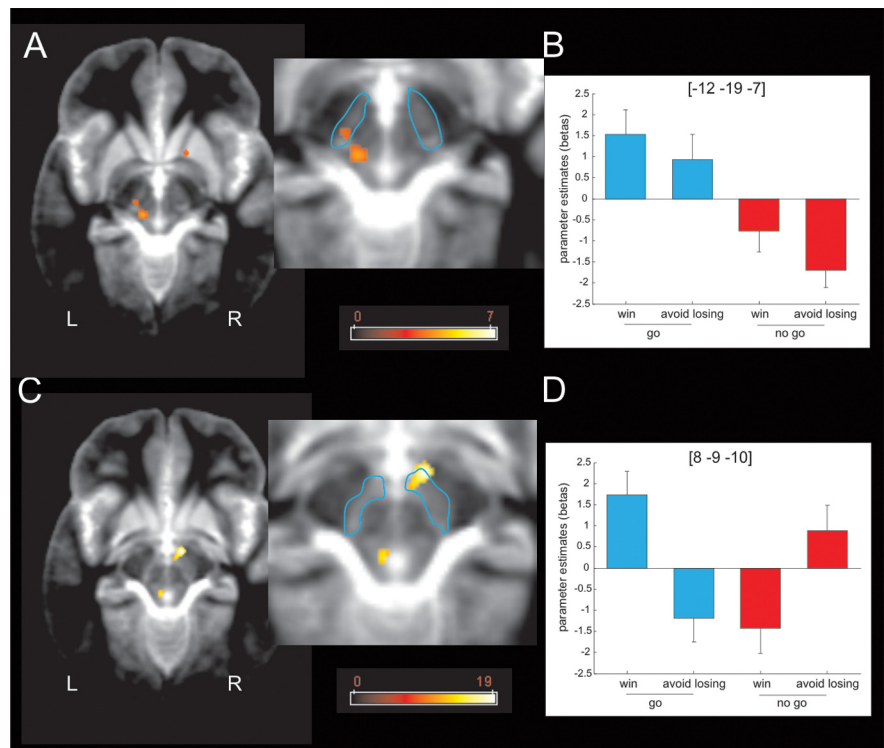
### Brain responses at outcome

Although our statistical model was suboptimal for studying brain responses at the time of the outcome, we performed a one-way ANOVA with valence as a factor (win, lose, or neutral) to confirm whether we could detect stronger BOLD responses in the ventral striatum for win than loss outcomes, an effect that has been described in many studies (for a recent review, see Haber and Knutson, 2010). As shown in Figure 6, our analysis revealed a simple main effect of valence in the right insula (whole-brain FWE,  $p < 0.05$ ), the left medial prefrontal cortex (whole-brain FWE,  $p < 0.05$ ), and ventral striatum (SVC,  $p < 0.05$ ). We did not find any activated voxels in SN/VTA. A *post hoc*  $t$  test analysis on peak voxels showed that the insula responded more to loss whereas ventral striatum responded more to wins, results broadly consistent with the existing literature (Haber and Knutson, 2010), findings that show that our imaging protocol was indeed sensitive to BOLD responses in the ventral striatum. Although our design was not optimal for studying outcome responses, this result demonstrates that the striatum responded to winning outcomes when consequences of an action were evaluated. This is in sharp contrast to the activation pattern seen during an anticipation period, which captured the influence of action requirements rather than valence. This pattern also fits well with the known role of striatum and dopaminergic system in reward-guided action learning (Robbins and Everitt, 2002; Frank et al., 2004).

### Discussion

Participants were faster and more successful in the go to win than the go to avoid losing condition. This suggests an asymmetric link between opponent response tendencies (go and no-go) and outcome valence (win and lose), consistent with a mandatory coupling between valence and action. Our parallel fMRI data showed that activation in striatum and lateral SN/VTA elicited by anticipatory cues predominantly represented a requirement for a go versus no-go response rather than the valence of the predicted outcome (Figs. 3, 4). Finally, activity in the medial SN/VTA mirrored the asymmetric link between action and valence (Fig. 5).

An essential backdrop to our results, and indeed the rationale for our experimental design, is the contrast between a seemingly ineluctable tie between valence and action spectra and their log-

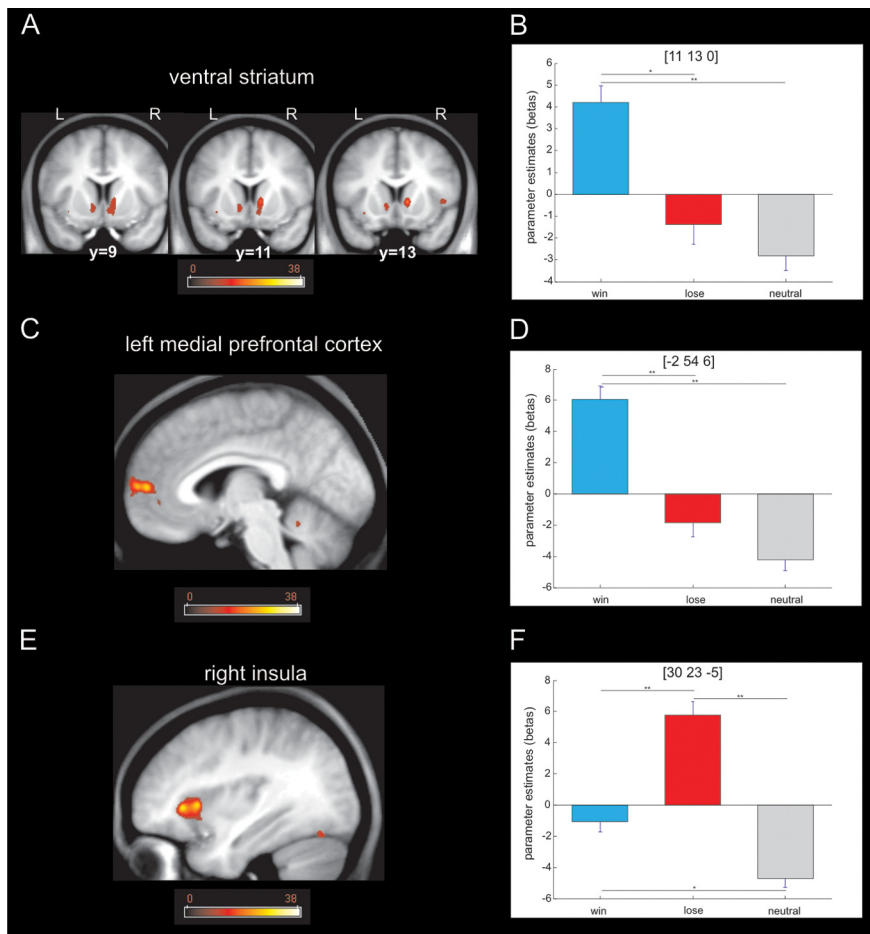


**Figure 5.** Midbrain response to anticipation of action and reward. **A**, Fractal images indicating that go trials elicited higher left lateral midbrain (or substantia nigra compacta) activity than fractal images indicating no-go trials ( $p < 0.001$  uncorrected;  $p < 0.05$  SVC). The color scale indicates  $t$  values. **B**, Parameter estimates at the peak coordinates in the left lateral midbrain confirm that activation at this location signals the anticipation of action regardless of the valence of the outcome of the action (reward or punishment avoidance). Coordinates are given in MNI space. Error bars indicate SEM (note that these parameter estimates were not used for statistical inference). **C**, An action  $\times$  valence interaction was observed in the right medial midbrain or ventral tegmental area ( $p < 0.001$  uncorrected;  $p < 0.05$  SVC). The color scale indicates  $F$  values. **D**, Parameter estimates at the peak coordinates in the right medial midbrain confirm that activation at this location signals the anticipation of action if the outcome of the action is rewarding. The anticipation of actions that avoid punishment, conversely, is associated with a relative deactivation of this region. The inverse pattern of activation is observed for the no-go trials: the anticipation of a passive response that wins a reward is associated with relative deactivation, whereas the anticipation of a passive response that avoids punishment is associated with activation. Coordinates are given in MNI space. Error bars indicate SEM (note that these parameter estimates were not used for statistical inference). L, Left; R, right.

ical independence. In particular, it is widely reported that dopamine neurons report a prediction error for reward (Montague et al., 1996; Schultz et al., 1997; Bayer and Glimcher, 2005) in the striatum (McClure et al., 2003; O'Doherty et al., 2003; O'Doherty et al., 2004). However, dopamine also invigorates action (Salamone et al., 2005), regardless of its instrumental appropriateness, with dopamine depletion being linked to decreased motor activity (Ungerstedt, 1971) and decreased vigor or motivation to work for rewards in demanding reinforcement schedules (Salamone and Correa, 2002; Niv et al., 2007). This coupling between action and reward in the dopaminergic system is exactly why a signal associated with go versus no-go might be confused with a signal associated with reward versus punishment.

### The role of action in a modified theory of opponency: striatum and lateral SN/VTA

Many previous fMRI experiments involving pavlovian and instrumental conditioning have reported BOLD signals in both striatum (McClure et al., 2003; O'Doherty et al., 2003, 2004) and SN/VTA (D'Ardenne et al., 2008), correlating with putative prediction error signals. In most studies [although not all, including those involving the anticipation of pain (Seymour et al., 2004, 2005) and monetary loss (Delgado et al., 2008)], these signals are



**Figure 6.** Brain responses to the outcome. **A**, Activation in the ventral striatum revealed by a one-way ANOVA with valence as factor ( $p < 0.001$  uncorrected;  $p < 0.05$  SVC). The color scale indicates  $F$  values. **B**, *Post hoc*  $t$  test on the peak voxel in ventral striatum revealed that the main effect was driven by higher activation for the win trials when compared with the loss and neutral trials ( $*p < 0.005$ ;  $**p < 0.001$ ). **C**, Activation in left medial prefrontal cortex revealed by a one-way ANOVA with valence as factor ( $p < 0.001$  uncorrected;  $p < 0.05$  whole-brain FWE). The color scale indicates  $F$  values. **D**, *Post hoc*  $t$  test on the peak voxel in left medial prefrontal cortex revealed a main effect driven by greater activation for win compared with loss and neutral trials, respectively ( $**p < 0.001$ ). **E**, Activation in right insula revealed by a one-way ANOVA with valence as factor ( $p < 0.001$  uncorrected;  $p < 0.05$  whole-brain FWE). The color scale indicates  $F$  values. **F**, *Post hoc*  $t$  test on the peak voxel in the right insula revealed that the main effect was driven by higher activation for loss trials compared with win and neutral trials ( $*p < 0.005$ ,  $**p < 0.001$ ). L, Left; R, right.

positive when the prediction of future gains is greater than expected and negative when the prediction of future losses is greater than expected. Our task was not designed to test the existence of such prediction errors (McClure et al., 2003; O'Doherty et al., 2003, 2004). Nevertheless, according to temporal difference learning, the prediction errors associated with the appearance of cues are the same as the predictions itself, and, given that performance of the task was indeed very stable throughout the period of fMRI data acquisition, we can reasonably assume that any brain region encoding a reward prediction error to presentation of cues might be expected to express a main effect of valence (go to win + no-go to win > go to avoid losing + no-go to avoid losing). Despite the clear effect of valence on behavioral performance, we did not find a main effect of valence in the fMRI data during anticipation, apart from a small cluster within left ventral putamen. Even there, cue-evoked activity for the go to win and no-go to win conditions differed significantly, despite both having the same expected value. We interpret these results as valence losing out to invigoration and thus as motivating the direct incorporation of action into theories of opponency.

action suppression activates inferior frontal gyrus (Rubia et al., 2003; Aron and Poldrack, 2006) and subthalamic nucleus (Aron and Poldrack, 2006). To our knowledge, suppression of neuronal responses in the striatum has not been systematically reported, although we note some evidence suggesting that striatal activity is enhanced by a need for action suppression (Aron et al., 2003; Aron and Poldrack, 2006). A second possibility arises from an alternative computational implementation for action choice in reinforcement learning. In the purest form of actor, the propensities to perform a given action are detached from the values of the states in which they are taken (Sutton and Barto, 1998). Thus, invigorating (or inhibiting) an action requires a positive (or negative) propensity for go, with the scale of the propensities being detached from any consideration of state values. Finally, a third possibility is that the striatum represents the advantage of making a go action as in advantage reinforcement learning (Dayan, 2002). In this model, action selection results from comparing the advantage of different options in which the advantage is the difference between the action value and the state value. Whereas state values would be positive in the win conditions and negative in the

In the one area of the ventral striatum in which we observed a main effect of valence, the BOLD response took a form that was more akin to the value (called a  $Q$  value) associated with the go action as opposed to one associated with a reward prediction or prediction error. That is, there was a single available action in our experiment, namely to generate a go response. The  $Q$  value of this response was high when an action was rewarded (go to win), zero when go responses led to avoidance of punishment (go to avoid losing), or omission of reward (no-go to win) and negative when actions were punished (no-go to avoid losing). The observation that the ventral subdivision of the striatum showed an action-dependent prediction was unexpected, given its association with the affective critic, which is governed by valence rather than some form of actor (O'Doherty et al., 2004). Although visual inspection of Figure 3 seems to suggest that this kind of signal is widely represented in most of our anatomical ROIs, especially the ventral subdivision of the striatum, statistical analyses do not support the presence of a systematic difference between go to win and go to avoid losing. However, we cannot entirely rule out the presence of such a signal.

The main effect in the striatum and lateral SN/VTA related most strongly to action (go to win + go to avoid losing > no-go to win + no-go to avoid losing). There are at least three possible interpretations for this dominance. First, it could be argued that the no-go condition requires inhibition of a prepotent motor response, and a relative deactivation in the striatum might reflect action suppression. However, there are good empirical grounds to believe this is not the case, including evidence from previous fMRI studies that

lose conditions, the advantage of performing a go action would be positive but small in the go conditions (because action values are positive or neutral) and negative in the no-go condition (because action value is neutral or negative). However, if the observed responses in the striatum and SN/VTA represented advantages, these would be the advantage of the go action even when participants successfully choose a no-go response.

Our results show that, during the anticipatory phase, striatal representations are dominated by actions rather than state values independent from action. These results are not incompatible with previous studies reporting reward prediction errors for state values under experimental conditions controlling action requirements indirectly through the use of explicit foil actions (Delgado et al., 2000, 2003, 2004; O'Doherty et al., 2003; Seymour et al., 2004; Tricomi et al., 2004). This is because, in those studies, reward prediction errors were isolated by comparing actions leading to rewards with foil actions that did not result in reward. Hence, those studies were suitable to isolate reward components that could be observed in addition to action representations in the striatum and SN/VTA. However, they were less suitable to highlight the predominant role of action representations in these regions. In fact, our results show that, when the action axis is explicitly incorporated within the experimental design, a refined picture of striatal and SN/VTA representations emerges. Our design allowed us to show that the predominant coding reflects anticipation of action. Reward prediction errors for state values may be superimposed either when an instrumental action is required to gain a reward or when an action tendency is automatically generated in response to a reward-predicting cue as in classical (pavlovian) conditioning.

In light of these results and within the limitation of fMRI studies of the SN/VTA (Düzel et al., 2009), theories implicating dopaminergic system in valence opponency (Daw et al., 2002) may need to be modified (Boureau and Dayan, 2011). The dopaminergic system would have to play a critical role in punishment as well as reward processing, whenever an action is required. That is, the semantics of the dopamine signal should be changed to reflect loss avoidance by action (Dayan and Huys, 2009) as well as the attainment of reward through action, indeed as in classical two-factor theories (Mowrer, 1947). In fact, some reinforcement learning models of active avoidance code the removal of the possibility of punishment (i.e., the achievement of safety) as akin (dopaminergically coded) to a reward (Grossberg, 1972; Schmajuk and Zanutto, 1997; Johnson et al., 2002; Moutoussis et al., 2008; Maia, 2010). Compatible with this two-factor view is the observation that dopamine depletion impairs the acquisition of active avoidance behavior (McCullough et al., 1993; Darvas et al., 2011). Paralleling the case for dopamine, this modification from valence opponency toward action opponency motivates a search for an identifiable neurotransmitter system that promotes the other end of the action spectrum, namely inhibition. Serotonin has been thought to serve as such a neurotransmitter (Deakin and Graeff, 1991; Gray and McNaughton, 2000). Interestingly, one study that inspired ours (Crockett et al., 2009) showed that tryptophan depletion abolished punishment-induced inhibition, which is similar to the disadvantage that we observed in the go to avoid losing condition.

### Medial SN/VTA

Unlike the case for the lateral SN/VTA, valence had opposite effects for go and no-go in the medial SN/VTA: for go, neural activity was higher for the win condition, whereas for no-go, activity was higher for the avoid-losing condition. One way to interpret this pattern is in terms of prediction errors relative to

the mandatory couplings between action and reward and between inhibition and punishment. That is, go is mandatorily associated with reward, and so the relevant prediction error, which could stamp in appropriate actions, favors reward over punishment. Conversely, no-go is associated with punishment, and so the relevant prediction error favors punishment over reward. Indeed, punishment prediction errors have been reported previously (Seymour et al., 2004; Delgado et al., 2008) in pavlovian conditions in which actions are irrelevant. Future studies could usefully target the functional interactions between action and valence in the medial SN/VTA, taking account also of anatomical and physiological findings regarding the involvement of dopamine in processing punishment. Unexpected punishment leads to supra-baseline dopamine activity in some microdialysis experiments in rats (Pezze et al., 2001; Young, 2004). Furthermore, unconditioned avoidance responses can only be elicited from topographically appropriate regions of the shell region of the nucleus accumbens given appropriately high levels of dopamine (Faure et al., 2008). Thus, one possibility is that the signal we observed in medial SN/VTA was more akin to one that organizes unconditioned responses in a valence-dependent manner.

### Conclusions

Our study expands on conventional views regarding the nature of signals reported from both STN/VTA and striatum. Although the striatum responded to wins more than losses at outcome, a primary form of coding in both the striatum and lateral SN/VTA complex during anticipation reflected action requirement rather than state values. These results indicate that the status of an action in relation to approach or withdrawal may be best captured in a modified opponent theory of dopamine function.

### References

- Aron AR, Poldrack RA (2006) Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *J Neurosci* 26:2424–2433.
- Aron AR, Schlaghecken F, Fletcher PC, Bullmore ET, Eimer M, Barker R, Sahakian BJ, Robbins TW (2003) Inhibition of subliminally primed responses is mediated by the caudate and thalamus: evidence from functional MRI and Huntington's disease. *Brain* 126:713–723.
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- 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.
- Boureau YL, Dayan P (2011) Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36:74–97.
- Brett M, Anton J-L, Valabregue R, Poline J-B (2002) Region of interest analysis using an SPM toolbox. Presented at the Eighth International Conference on Functional Mapping of the Human Brain, Sendai, Japan, June.
- Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci U S A* 106:4894–4899.
- Bunzeck N, Düzel E (2006) Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51:369–379.
- Cools R, Nakamura K, Daw ND (2011) Serotonin and dopamine: unifying affective, motivational, and decision functions. *Neuropsychopharmacology* 36:98–113.
- Cooper JC, Knutson B (2008) Valence and salience contribute to nucleus accumbens activation. *Neuroimage* 39:538–547.
- 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.
- 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.



- Darvas M, Fadok JP, Palmiter RD (2011) Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. *Learn Mem* 18:136–143.
- Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616.
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Dayan P (2002) Motivated reinforcement learning. In: *Advances in neural information processing systems 14* (Dietterich TG, Becker S, Ghahramani Z, eds). Cambridge, MA: MIT.
- Dayan P, Huys QJ (2009) Serotonin in affective control. *Annu Rev Neurosci* 32:95–126.
- Deakin JFW, Graeff FG (1991) 5-HT and mechanisms of defence. *J Psychopharmacol* 5:305–315.
- 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.
- Delgado MR, Locke HM, Stenger VA, Fiez JA (2003) Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn Affect Behav Neurosci* 3:27–38.
- Delgado MR, Stenger VA, Fiez JA (2004) Motivation-dependent responses in the human caudate nucleus. *Cereb Cortex* 14:1022–1030.
- Delgado MR, Li J, Schiller D, Phelps EA (2008) The role of the striatum in aversive learning and aversive prediction errors. *Philos Trans R Soc Lond B Biol Sci* 363:3787–3800.
- Domjan M (2005) Pavlovian conditioning: a functional perspective. *Annu Rev Psychol* 56:179–206.
- Düzel E, Bunzeck N, Guitart-Masip M, Wittmann B, Schott BH, Tobler PN (2009) Functional imaging of the human dopaminergic midbrain. *Trends Neurosci* 32:321–328.
- 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.
- Faure A, Reynolds SM, Richard JM, Berridge KC (2008) Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. *J Neurosci* 28:7184–7192.
- Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Gerfen CR (1992) The neostriatal mosaic: multiple levels of compartmental organization. *Trends Neurosci* 15:133–139.
- Gray JA, McNaughton M (2000) *The neuropsychology of anxiety: an inquiry into the function of the septohippocampal system*, Ed 2. Oxford: Oxford UP.
- Grossberg S (1972) A neural theory of punishment and avoidance. *Math Biosci* 15:39–67.
- Guitart-Masip M, Bunzeck N, Stephan KE, Dolan RJ, Düzel E (2010) Contextual novelty changes reward representations in the striatum. *J Neurosci* 30:1721–1726.
- Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35:4–26.
- Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 20:2369–2382.
- Johnson J, Li W, Li J, Klopff A (2002) A computational model of learned avoidance behavior in a one-way avoidance experiment. *Adapt Behav* 9:91–104.
- Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, Christensen GE, Collins DL, Gee J, Hellier P, Song JH, Jenkinson M, Lepage C, Rueckert D, Thompson P, Vercauteren T, Woods RP, Mann JJ, Parsey RV (2009) Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46:786–802.
- Knutson B, Adams CM, Fong GW, Hommer D (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159(1–5).
- Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005) Distributed neural representation of expected value. *J Neurosci* 25:4806–4812.
- Lammel S, Hetzel A, Häckel O, Jones I, Liss B, Roeper J (2008) Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* 57:760–773.
- Maia TV (2010) Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn Behav* 38:50–67.
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- McCullough LD, Sokolowski JD, Salamone JD (1993) A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience* 52:919–925.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Moutoussis M, Bental RP, Williams J, Dayan P (2008) A temporal difference account of avoidance learning. *Network* 19:137–160.
- Mowrer OH (1947) On the dual nature of learning: a reinterpretation of conditioning and problem solving. *Harv Educ Rev* 17:102–148.
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.
- O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.
- O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. *Neuron* 33:815–826.
- 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.
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454.
- Pezze MA, Heidbreder CA, Feldon J, Murphy CA (2001) Selective responding of nucleus accumbens core and shell dopamine to aversively conditioned contextual and discrete stimuli. *Neuroscience* 108:91–102.
- Redgrave P, Prescott TJ, Gurney K (1999) Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 22:146–151.
- Robbins TW, Everitt BJ (2002) Limbic-striatal memory systems and drug addiction. *Neurobiol Learn Mem* 78:625–636.
- Rubia K, Smith AB, Brammer MJ, Taylor E (2003) Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage* 20:351–358.
- Salamone JD, Correa M (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav Brain Res* 137:3–25.
- 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.
- Schmajuk N, Zanotto B (1997) Escape avoidance and imitation: A neural network approach. *Adapt Behav* 6:63–129.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667.
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci* 8:1234–1240.
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. Cambridge, MA: Massachusetts Institute of Technology.
- Tricomi EM, Delgado MR, Fiez JA (2004) Modulation of caudate activity by action contingency. *Neuron* 41:281–292.
- Ungerstedt U (1971) Adipsia and aphagia after 6-hydroxydopamine induced degeneration of the nigro-striatal dopamine system. *Acta Physiol Scand Suppl* 367:95–122.
- Weiskopf N, Hutton C, Josephs O, Deichmann R (2006) Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. *Neuroimage* 33:493–504.
- Wrase J, Kahnt T, Schlagenhauf F, Beck A, Cohen MX, Knutson B, Heinz A (2007) Different neural systems adjust motor behavior in response to reward and punishment. *Neuroimage* 36:1253–1262.
- Young AM (2004) Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1 min microdialysis in rats. *J Neurosci Methods* 138:57–63.
- Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS (2003) Human striatal response to salient nonrewarding stimuli. *J Neurosci* 23:8092–8097.
- Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004) Human striatal responses to monetary reward depend on saliency. *Neuron* 42:509–517.