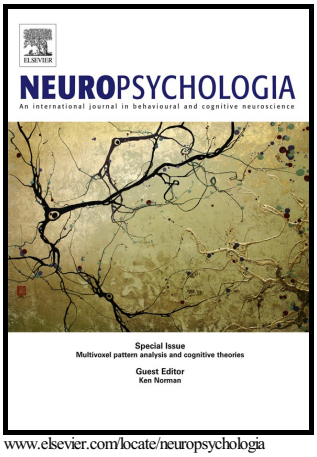

www.elsevier.com/locate/neuropsychologia

The left hemisphere learns what is right: Hemispatial reward learning depends on reinforcement learning processes in the contralateral hemisphere

Kristoffer Carl Aberg, Kimberly Crystal Doell, Sophie Schwartz

PII: S0028-3932(16)30174-9
DOI: http://dx.doi.org/10.1016/j.neuropsychologia.2016.05.023
Reference: NSY6004

To appear in: *Neuropsychologia*

Received date: 14 January 2016
Revised date: 19 April 2016
Accepted date: 21 May 2016

Cite this article as: Kristoffer Carl Aberg, Kimberly Crystal Doell and Sophie Schwartz, The left hemisphere learns what is right: Hemispatial reward learning depends on reinforcement learning processes in the contralateral hemisphere *Neuropsychologia,* http://dx.doi.org/10.1016/j.neuropsychologia.2016.05.023

Title: The left hemisphere learns what is right: Hemispatial reward learning depends on reinforcement learning processes in the contralateral hemisphere

Authors: Kristoffer Carl Aberg[1,2,3] Kimberly Crystal Doell[1,2,3] Sophie Schwartz[1,2,3]

1. Department of Neuroscience, Faculty of Medicine, University of Geneva, Switzerland

2. Swiss Center for Affective Sciences, University of Geneva, Switzerland

3. Geneva Neuroscience Center, University of Geneva, Switzerland

Corresponding author: Kristoffer Carl Aberg

University Medical Center, CMU
Bat. B, Dept of Neuroscience, 7th floor, room 7004
1 rue Michel-Servet, CH-1211 GENEVA 4

Tel : +41 (0)22 379 53 61

Fax : +41 (0)22 379 54 02

Email: kc.aberg@gmail.com

Abstract

Orienting biases refer to consistent, trait-like direction of attention or locomotion toward one side of space. Recent studies suggest that such hemispatial biases may determine how well people memorize information presented in the left or right hemifield. Moreover, lesion studies indicate that learning rewarded stimuli in one hemispace depends on the integrity of the contralateral striatum. However, the exact neural and computational mechanisms underlying the influence of individual orienting biases on reward learning remain unclear. Because reward-based behavioral adaptation depends on the dopaminergic system and prediction error (PE) encoding in the ventral striatum, we hypothesized that hemispheric asymmetries in dopamine (DA) function may determine individual spatial biases in reward learning. To test this prediction, we acquired fMRI in 33 healthy human

participants while they performed a lateralized reward task. Learning differences between hemispaces were assessed by presenting stimuli, assigned to different reward probabilities, to the left or right of central fixation, i.e. presented in the left or right visual hemifield. Hemispheric differences in DA function were estimated through differential fMRI responses to positive vs. negative feedback in the left vs. right ventral striatum, and a computational approach was used to identify the neural correlates of PEs. Our results show that spatial biases favoring reward learning in the right (vs. left) hemifield were associated with increased reward responses in the left hemisphere and relatively better neural encoding of PEs for stimuli presented in the right (vs. left) hemifield. These findings demonstrate that trait-like spatial biases implicate hemisphere-specific learning mechanisms, with individual differences between hemispheres contributing to reinforcing spatial biases.

Keywords

Orienting; Reward; Reinforcement learning; Prediction error; Hemisphere; Dopamine

1. Introduction

Spatial orienting biases implicate a preferred and consistent direction of attention or locomotion toward one side of space (i.e. hemispace). Such biases can be expressed as increased body rotations toward one side (Mohr, Brugger, Bracha, Landis, & Viaud-Delmon, 2004) or better detection of targets located in one hemispace (Nash, McGregor, & Inzlicht, 2010; Tomer, et al., 2013). Orienting biases have been attributed to the imbalanced activation of hemispheres (Kinsbourne, 1970; Nash, et al., 2010), with, for example, increased orienting responses to stimuli presented to the hemifield contralateral to the more activated hemisphere (Kelly, Gomez-Ramirez, & Foxe, 2009; Newman, O'Connell, & Bellgrove, 2013; Thut, Nietzel, Brandt, & Pascual-Leone, 2006). Moreover, orienting

biases have been found to be consistent over time in both humans and animals, therefore indicating a trait-like disposition (Andrade, Alwarshetty, Sudha, & Suresh Chandra, 2001; Tomer, 2008).

There is evidence suggesting that traits influence individual learning processes. For example, participants displaying high trait optimism show strong deficits in learning information that is worse than expected, i.e. information inconsistent with their optimistic predisposition (Sharot, Korn, & Dolan, 2011). Conversely, trait anxiety impedes fear extinction (Indovina, Robbins, Nunez-Elizalde, Dunn, & Bishop, 2011), and high sensitivity to social rejection prevents extinction of conditioned responses to angry faces (Olsson, Carmona, Downey, Bolger, & Ochsner, 2013). Thus, trait-like dispositions may influence learning mechanisms so as to reinforce and maintain trait-related behavioural biases. Similarly, animal data suggest that orienting biases could affect spatial learning. Rats displaying preferences for one arm in a T-maze showed strong deficits when learning to enter the arm contralateral to their bias, but not when learning to enter the ipsilateral arm (Andrade, et al., 2001). However, it is not known whether individual orienting biases may also influence hemispatial learning in humans. More generally, the computational and neural learning mechanisms underlying trait-related effects on learning are largely understudied.

On the one hand, both animals and humans orient away from the hemisphere with more dopamine (DA) activity, as found in both locomotion (Barone, Bankiewicz, Corsini, Kopin, & Chase, 1987; Bracha, Shults, Glick, & Kleinman, 1987) and attention (Lee, Harris, Atkinson, & Fowler, 2001; Tomer, et al., 2013). Given that orienting biases are believed to be caused by imbalanced activity between hemispheres (Kinsbourne, 1970), together with activational aspects of DA (Robbins & Everitt, 1992, 2007), orienting biases caused by the asymmetric activation of hemispheres could be mediated by asymmetries in DA function (Tomer, et al., 2013). For example, it is plausible to assume that orientation is biased towards visual stimuli presented in the hemifield contralateral to the hemisphere with relatively increased DA function. On the other hand, there is evidence suggesting that the striatum, a subcortical brain region receiving projections from midbrain DA neurons

(Glimcher, 2011), is important for hemispatial reward learning. For example, Lucas et al. (2013) investigated hemispatial reward learning in stroke patients displaying hemispatial neglect. In their experiment, patients gained more rewards for detecting targets in the neglected hemispace, leading to transient reductions of hemispatial neglect. Importantly, some patients displayed hemispatial learning deficits, and careful examination revealed that their lesions extended into the contralateral (i.e. ipsilesional) striatum. Similarly, patients with lesions extending into the contralateral striatum were unable to show reward-related reductions of hemispatial neglect (Malhotra, Soto, Li, & Russell, 2013). Together these results suggest that hemispheric asymmetries in striatal function may not only subtend orienting biases, but could also be involved in biasing hemispatial learning. Because DA levels directly influence the neural encoding of prediction errors (PEs), i.e. the mismatch between actual and predicted outcomes (Chowdhury, et al., 2013; Jocham, Klein, & Ullsperger, 2011; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Sutton & Barto, 1998) in the striatum (Abler, Walter, Erk, Kammerer, & Spitzer, 2006; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003), here we hypothesized that hemispheric asymmetries in DA function could be associated with hemispheric asymmetries in the neural encoding of PEs. In particular, a relative increase of DA function in one hemisphere would lead to better encoding of PEs related to stimuli presented in the contralateral (vs. ipsilateral) hemispace, and thus also result in hemispatial learning biases. Supporting this hypothesis, one recent study applied deep-brain stimulation of the subthalamic nucleus in one hemisphere, a procedure known to mimic DA enhancers, and found enhanced reward-learning for stimuli presented to the contralateral visual hemifield (Palminteri, et al., 2013). However, the neural structures (beyond the subthalamic nucleus) involved in mediating this effect were not investigated.

In the present study, we tested whether an asymmetry in neural activity across dopaminergic regions in favor of the left (vs. right) hemisphere implicates better reward learning in the right (vs. left) hemispace, and how such hemispatial biases in reward learning may be explained in terms of underlying computational and neuronal learning mechanisms. Specifically, we investigated individual biases in reward learning between hemispaces by presenting stimuli with different reward

probabilities to the left and right of a central fixation, i.e. in the left and right hemifields (Frank, Seeberger, & O'Reilly, 2004). Based on the close connection between DA function and reward processing in the striatum (Knutson & Gibbs, 2007; van der Vegt, et al., 2013), we estimated hemispheric differences in DA function during the probabilistic selection task as the differential neural response of the left vs. right ventral striatum to positive vs. negative feedback (Aberg, Doell, & Schwartz, 2015). We then used a computational reinforcement learning model to identify modulations of fMRI signal corresponding to the PE for stimuli presented in each hemifield.

The results show that participants displaying a larger reward response in the left (vs. the right) ventral striatum learned rewarded stimuli better when they were presented in the right (vs. the left) hemifield, as indexed by higher learning rates. Moreover, these participants encoded PEs better in the left ventral striatum for stimuli presented in the right (vs. the left) hemifield. These data uncover a neural mechanism whereby the expression of hemispatial biases during learning implicates hemispheric asymmetries in DA-related functions.

2. Materials and Methods

2.1 Participants

After having provided written consent according to the ethical regulations of the Geneva University Hospital, 42 healthy volunteers participated in this experiment which was conducted in accordance with the Declaration of Helsinki. All participants were right-handed, native French speakers, and without any previous history of neurological or psychological disorders. Two participants failed to follow task instructions, four did not reach the performance criteria (see below), two fell asleep in the MRI scanner, and the data of one participant was excluded because of a bad fit between the behaviour data and the computational model. Finally, data from 33 participants [18 females; age=23.61 ± 0.79 years ± SEM] were included in the analyses.

2.2 Probabilistic selection task (PST)

In the PST, reward probabilities associated with different symbols were learned through a trial-and-error procedure (Frank, Woroch, & Curran, 2005). During the (main) training period, taking place outside the MRI scanner, a central fixation cross (average presentation time: 1.25 s; jitter range: 0.5-2.0 s) signalled the start of a trial, followed by a pair of symbols (presentation time: 1.0 s) presented to the left and to the right of fixation (Figure 1A). Selection of a symbol (left or right) was done by pressing its corresponding button (left or right) with the right hand and resulted in the presentation of either a happy or a sad smiley face, i.e. positive and negative feedback (presentation time: 0.6 s). The feedback type depended on a reward probability associated with the selected symbol (Figure 1B), i.e. selection of the A symbol yielded positive feedback in 80% of the trials while selecting the B symbol resulted in negative feedback in 80% of the trials. Participants were instructed to collect as many happy smiley faces as possible while avoiding getting sad smiley faces. Each of the three pairs of symbols (i.e. AB, CD, and EF pairs) were presented 20 times in blocks of 60 trials in a pseudorandom order. This ensured that one pair was not repeated until one trial of each other pair had been presented. Critically, each symbol was presented an equal number of times to each hemifield. Between participants, the symbols were randomly assigned to different pairs and reward probabilities were randomly assigned to different symbols. To guarantee that some learning had occurred before entering the MRI scanner, participants had to reach a selection rate of the A and C symbols of more than 60% (12/20 trials) and 55% (11/20 trials), respectively, in one block of 60 trials, while no criteria were enforced for the more difficult EF pair (for a similar procedure see Frank et al., 2005). Four participants did not reach the criteria within 45 minutes of training, and their data were not included in the analyses.

Two additional training blocks of 60 trials were then performed inside the MRI scanner in order to investigate the neural correlates of reward processing and reinforcement learning. To make the task design suitable for an event-related fMRI design, the average presentation time for the central fixation cross was changed to 3.5 s with a jitter range between 2.0 and 5.0 s.

During the subsequent test phase, twelve additional novel pairs (AC, AD, AE, AF, BC, BD, BE, BF, CE, CF, DE, and DF) were created by combining the symbols from the original trained pairs (AB, CD, and EF). This phase was largely similar to the training phase inside the scanner except that feedback was no longer provided in order to prevent further learning of the new pairs. Thus, this phase tested the generalization of learned symbol values from the trained pairs to the novel pairs, and participants were instructed to perform the task as well as possible and to trust their instincts, or guess, when uncertain. Each of the 15 pairs were presented eight times in a pseudorandom order with each pair presented once before any other pair was repeated, and each symbol was presented an equal number of times to each hemifield. The fMRI data analyses are explained in the MRI Data Analysis section below (e.g. Section 2.5).

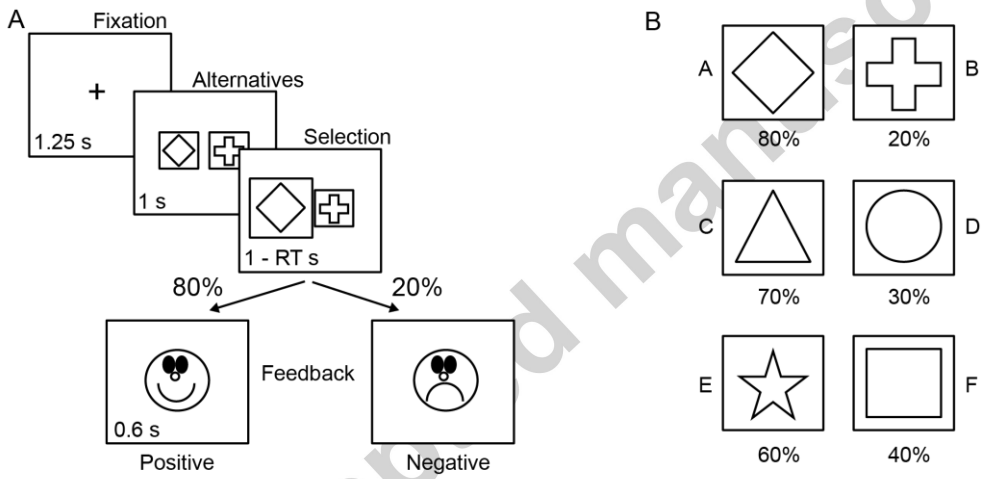

Figure 1. A. A training trial in the probabilistic selection task. After fixation, one symbol was presented to the left hemifield and another symbol was presented to the right hemifield. Participants had to select one symbol within 1s. A happy or a sad smiley face was then presented based on the reward probability associated with the selected symbol. RT = response time. B. Reward probabilities associated with each pair and symbol. The symbols associated with each reward probability were randomized between participants.

2.2.1 Asymmetry indexes

This study investigates how hemispatial biases expressed after learning relate to asymmetries in reward and reinforcement learning processes. As subtle differences in reward learning are more likely to be detected for the most rewarded stimuli, subsequent analyses will focus on the A symbol

with the highest reward probability (Frank, et al., 2004). Specifically, the presence of post-learning hemispatial biases is estimated during the test phase as the difference in selection rates for the most rewarded A symbol when presented to the right (vs. the left) hemifield.

$$\text{Selection asymmetry Index} = \frac{\%\text{Selected}_{\text{A,RVF}} - \%\text{Selected}_{\text{A,LVF}}}{\%\text{Selected}_{\text{A,RVF}} + \%\text{Selected}_{\text{A,LVF}}}$$

A positive selection asymmetry index indicates that the selection rate of the A symbol is higher when presented in the right hemifield (RHF) vs. left hemifield (LHF), indicating better reward learning in the RHF.

However, better reward learning in one hemifield could also lead to increased incentive motivation for highly rewarded objects in this hemifield (Palminteri, et al., 2013; Schmidt, Palminteri, Lafargue, & Pessiglione, 2010). Incentive motivation is associated with increased response vigor (Niv, Daw, Joel, & Dayan, 2007) and reduced response times (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Knutson, Fong, Adams, Varner, & Hommer, 2001). For this purpose, a response time (RT) asymmetry index was calculated as another measure of post-learning, reward-related hemispatial biases.

$$\text{RT asymmetry Index} = \frac{\text{RT}_{\text{A,LVF}} - \text{RT}_{\text{A,RVF}}}{\text{RT}_{\text{A,LVF}} + \text{RT}_{\text{A,RVF}}}$$

A positive RT asymmetry index indicates that the RT of the A symbol was faster in the RHF (vs. the LHF), indicating better reward learning in the RHF. As the RT provides a continuous estimate of incentive motivation, it may more accurately reflect relative differences in incentive motivation between hemifields, as compared to selection rates which are related to discrete reward probabilities.

2.3 Reinforcement learning model

Computational learning models provide a mechanistic approach to study learning-related trial-by-trial variations in behaviour (Watkins & Dayan, 1992) and in neural activity (Glascher, Daw, Dayan,

& O'Doherty, 2010). Moreover, such models can be used to summarize complex learning behaviours into a few parameters which can be used to link differences in post-learning behaviour to differences in learning processes (Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). Here, we used a modified version of the classical Q-learning model to account for differences in learning from different hemifields (for similar approaches, see Palminteri, Boraud, Lafargue, Dubois, & Pessiglione, 2009; Palminteri, et al., 2013). In this modified Q-learning model, each symbol is assigned a value $Q_{i,j}$ which depends on its reward history. The value of $Q_{i,j}$ is updated each time the corresponding symbol i has been selected in the hemifield j:

$$Q_{i,j}(t+1) = Q_{i,j}(t) + \alpha_j(r - Q_{i,j}(t))$$

where α_j is the learning rate for items selected in the hemifield j and r is the reward outcome (set to 1 for positive outcomes and 0 for negative outcomes). The probability of selecting a specific symbol is modeled by a softmax rule:

$$p_{A,LHF}(t) = \frac{e^{\frac{Q_{A,LHF}(t)}{\beta}}}{e^{\frac{Q_{A,LHF}(t)}{\beta}} + e^{\frac{Q_{B,RHF}(t)}{\beta}}}$$

where in this example, $p_{A,LHF}$ is the probability of selecting symbol A presented to the left hemifield in an AB pair. The β controls 'exploit vs. explore' behavior. When this parameter is small, symbols with the highest Q value are most likely selected (exploitation) while a large β leads to selections less dependent on the symbol's value (exploration). The three parameters α_{LHF},α_{RHF}, and β are fit to each participant's training data by minimizing the negative log likelihood estimate (LLE):

$$LLE = -\log(\prod_{1}^{n} p_{i,j}(t))$$

where $p_{i,j}(t)$ is the probability of selecting symbol i presented to the hemifield j in trial t.

Reinforcement learning crucially depends on the feedback, as no updating of Q-values (i.e. learning) occurs in the absence of feedback. Therefore, it is assumed that the Q-values obtained at the end of training remained constant throughout the test phase.

2.4 Statistical methods

ANOVAs were used to estimate statistical significance between more than two conditions. However, in small sample sizes Monte-Carlo randomization methods are particularly useful when estimating statistical significance between two conditions as these methods do not require any assumption concerning the distribution of data (Howell, 2013). Accordingly, we estimated statistical significance between two conditions by comparing original observed effects with a null-distribution created from 10000 samples, each created by shuffling the original data between conditions.

2.5 MRI data

2.5.1 Image acquisition

MRI images were acquired using a 3T whole body MRI scanner (Trio TIM, Siemens, Germany) with a 12-channel head coil. Functional images were acquired with a susceptibility weighted EPI sequence (voxel dimensions = 3.2 mm isotropic, TR/TE = 2100/30 ms, 64 x 64 x 36 voxels, flip angle = 80 degrees). Standard structural images were acquired with a T1 weighted 3D sequence (MPRAGE, voxel dimensions = 1 mm isotropic, TR/TI/TE = 1900/900/2.27 ms, 256 x 256 x 192 voxels, flip angle = 9 degrees). Proton density (PD) structural images were acquired with a turbo spin echo sequence (voxel dimensions = 0.8x0.8x3 mm, TR/TE = 6000/8.4 ms, 205 x 205 x 60 voxels, flip angle = 149 degrees). The PD scan was added in order to localize dopaminergic brain regions in the midbrain, i.e. the ventral tegmental area/substantia nigra (VTA/SN). For this reason, the acquisition volume was oriented in order to scan the brain from the top of the thalamus to the lower part of the pons.

2.5.2 MRI Data Analysis

All fMRI data were pre-processed and then analysed using the general linear model (GLM) for event-related designs in SPM8 (Welcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm). During pre-processing, all functional volumes were realigned to the mean image, co-registered to the structural T1 image, corrected for slice timing, normalized to the MNI EPI-template, and smoothed using an 8 mm FWHM Gaussian kernel. Statistical analyses

were performed on a voxelwise basis across the whole-brain. At the first-level analysis, individual events were modelled by a standard synthetic hemodynamic response function (HRF) and six rigid-body realignment parameters were included as nuisance covariates when estimating statistical maps. Contrasts between conditions (see below) were then calculated and the contrast images entered into second-level tests implemented in SPM.

2.5.2.1 Reward processing and reinforcement learning

We assessed the neural correlates of prediction errors (PEs) and hemispheric asymmetries in reward processing from the data obtained during the two training blocks performed inside the MRI scanner. To investigate the neural correlates of reward processing, we first created an event-related design with two event-types time-locked to the onset of positive and negative feedback. We computed the difference in brain activity between positive and negative feedback for the left vs. the right ventral striatum to index the degree of hemispheric reward asymmetry (see Aberg, et al., 2015).

$$\text{Hemispheric reward asymmetry} = [\text{Pos.} - \text{Neg. FB}]_{\text{R VStr}} - [\text{Pos.} - \text{Neg. FB}]_{\text{L VStr}}$$

To investigate the neural correlates of PEs, we generated another model with two event-types time-locked to the onset of feedback after selecting symbols in the RHF and the LHF. Model-derived PEs related to each HF (i.e. PE_{RHF} and PE_{LHF}) were added as parametric modulators for each respective event-type. By measuring how strongly PEs corresponding to each hemifield modulated activity in the left and the right ventral striatum (see Section 2.5.3), we could test whether PEs related to one hemifield were better represented in the contralateral hemisphere. We also assessed PE encoding within dopaminergic midbrain regions (see Section 2.5.3), because these are known to contribute to PE encoding (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Schultz & Dickinson, 2000; Tobler, Fiorillo, & Schultz, 2005), which itself was recently shown to relate to the expression of post-learning behavioural biases (Aberg, et al., 2015).

2.5.2.1 Incentive values during testing

The neural correlates of incentive values were investigated during the test phase inside the MRI scanner. To assess differences in the neural response when the A symbol was presented to different hemifields, we used an event-related design with 24 event-types (12 different novel pairs and two hemifields; AC, CA, AD, DA … DE, ED, EF, FE) time-locked to the onset of stimulus presentation. The contrast for each symbol and hemifield was obtained by the linear combination of all trials containing the symbol presented to the corresponding hemifield (e.g. A_{RHF} = CA, DA, EA, FA). To specifically investigate neural responses related to hemifield-differences in incentive value within a trial, we created another event-related design with one event-type time-locked to the onset of stimulus presentation, with the relative values between symbols presented to different hemifields added as a parametric modulator for each trial (i.e. Q_{RHF}-Q_{LHF}). The ventral striatum and the dopaminergic midbrain have been implicated in incentive motivation and the representation of incentive values (Bartra, McGuire, & Kable, 2013; Knutson, Fong, et al., 2001; Knutson & Gibbs, 2007; Knutson, Taylor, Kaufman, Peterson, & Glover, 2005). Thus, to test for learning-induced biases in incentive motivation, brain activity evoked by value-differences between symbols presented to different hemifields (i.e. Q_{RHF}-Q_{LHF} and A_{RHF}-A_{LHF}) was investigated within these brain regions (see Section 2.5.3).

2.5.3 Region of interest (ROI)

A priori ROIs used for small volume corrections (SVC) and extraction of beta parameter estimates were created based on previous literature. For the ventral striatum ROIs, centre coordinates were obtained from a recent study [Neta, Oliveira, Correia, & Ferreira, 2008; left ventral striatum: MNI x=-9 y=9 z=-8; right ventral striatum: MNI x=9 y=8 z=-8]. The left and right ventral striatum ROIs were defined as spheres (radius=5mm) centred on these coordinates. For the SVC specifically, a bilateral ventral striatum mask was created by combining these two spheres into one mask. For the dopaminergic midbrain (i.e. the VTA) ROI, centre coordinates were obtained from a recent study [Ballard et al., 2011; left VTA x=-2 y=-15 z=-13; right VTA x=3 y=-17 z=-12]. The VTA ROI was created by combining two spheres (radius=4 mm) into one VTA mask, in accordance with a recently described procedure (Cha, et al., 2014).

2.5.4 Statistical Analyses

For the ventral striatum, the obtained results are reported using a threshold of $p<0.001$ and a minimum cluster size of five contiguous voxels. For the VTA, the obtained results are reported using a threshold of $p<0.005$ and a minimum cluster size of ten contiguous voxels (see Cha, et al., 2014). Small volume corrections (SVC) using a threshold of $p<0.05$ Family-Wise Error Rate for multiple comparisons were obtained using the a priori ROIs described above. Conditions were compared using traditional t-tests implemented in SPM. Due to recent concerns (see Vul, Harris, Winkielman, & Pashler, 2009), displayed effect sizes were calculated by averaging extracted beta values over all voxels within a ROI.

3. Results

3.1 Behavioral results

First, we used a linear regression analysis to ensure that training increased performance. Next, we tested whether the learned reward-values (derived from the training phase) influenced performance in the test phase, by comparing the selection rates and reaction times between symbols in the novel pairs, with a particular focus on performance differences between hemifields for the highly rewarded A-symbol (see Section 2.2.1 above). Finally, to unveil computational biases underpinning reward-related behavioural biases, we compared the model-derived results from the training phase with the behaviour during the test phase.

3.1.1 Training

Figure 2A displays performance for the different pairs as a function of training. Because participants needed a different numbers of trials to reach the performance criteria, performance was averaged across ten bins of trials for each participant and for each pair i.e. the bin size for a participant requiring 120 trials (i.e. 2 blocks of training) is 4 trials per bin while the bin size is 6 trials for a participant requiring 180 trials (e.g. 3 blocks of training). On average, a bin consisted of

7.47±1.11 (SEM) trials, indicating that participants on average required between 3-4 blocks of training to reach the criteria.

As a measure of learning, performance was linearly regressed over trial bins for each pair and participant, and the resulting regression coefficients (indicative of learning speed) were submitted to an ANOVA with factor Pair (AB, CD, EF). There was a significant effect of Pair [$F(2,64)=4.442$, $p=0.016$] because the regression coefficients were significantly larger for AB pairs [regression coefficient=0.022±0.004 (SEM)] as compared to EF pairs [regression coefficient=0.005±0.004 (SEM), $p=0.007$], and marginally so for CD pairs [regression coefficient=0.016±0.004 (SEM)] as compared to EF pairs [$p=0.061$]. Regression coefficients did not differ between AB and CD pairs [$p=0.335$]. Additionally, the average regressions coefficients were significantly larger than 0.0 for both AB [$p<0.0001$] and CD [$p=0.0005$] pairs, but not EF pairs [$p=0.170$]. These results show that learning occurred and was easier in pairs where the difference in reward probability between the symbols was large.

3.1.2 Testing

3.1.2.1 Selection rates

On average the frequently rewarded A symbol was not selected more often in the right hemifield (RHF) over the left hemifield (LVF) [Selection rate A: RHF=0.839±0.029 (SEM), LVF=0.826±0.032 (SEM), $p = 0.526$] (Fig 2B). However, and as expected from previous work (Tomer, 2008; Tomer, et al., 2013), there was a large individual variability: 11 and 15 participants selected the A symbol more frequently in the LHF and the RHF, respectively, while 7 participants displayed no spatial selection bias. Next, to ensure that participants were able to generalize learned symbol values to the novel pairs during testing, and to test whether there was any general selection bias favouring one hemifield, a repeated measures ANOVA with factors Symbol (A, B, C, D, E, F) and Side (LHF, RHF) was conducted. There was a main effect of Symbol [$F(5,352)=14.181$, $p<0.0001$] because symbols associated with high reward probabilities during training were more likely to be selected during

testing with novel pairs. This result suggests that the learned reward probabilities generalized and guided decision making also for the novel pairs. There was no significant effect of Side or interaction between Side and Symbol [both p-values>0.335].

3.1.2.2 Response times

Participants responded on average quicker when the A symbol was presented to the RHF as compared to the LHF [RT A: RHF=598.915±11.891 (SEM), LHF=616.009±11.290 (SEM), p=0.002] (Fig 2C), with 23 and 10 participants responding quicker when the A symbol was presented to the RHF and the LHF, respectively. To determine whether RTs were generally faster for symbols presented to one hemifield, a repeated measures ANOVA with factors Symbol (A, B, C, D, E, F) and Side (LVF, RVF) was calculated. There was a main effect of Symbol [$F(5,352)=84.689$, $p<0.0001$] because RTs were faster for symbols associated with higher reward probabilities during training. This result is in line with previous studies showing that faster RTs are associated with increased motivational incentives and reward values (Adcock, et al., 2006; Knutson, Fong, et al., 2001; Mir, et al., 2011). There was no significant main effect of Side or interaction between Side and Symbol [both p-values>0.665].

According to the theory of spatial priority maps, performance differences between the RHF and the LHF could be due to more general differences between hemifields which are independent of the incentive values of presented stimuli, i.e. that stimuli presented to one hemifield have higher priority than those presented to the other hemifield (Chelazzi, et al., 2014; Itti & Koch, 2001). To test whether this is the case in the present study, we tested whether individual differences in the hemifield bias expressed for the A symbol (i.e. the RT asymmetry index) correlated with similar biases for the other symbols (i.e. B_{LHF}-B_{RHF}, C_{LHF}-C_{RHF}, …, F_{LHF}-F_{RHF}). There were no significant correlations [all uncorrected p-values > 0.05], suggesting that the expression of individual spatial biases can be attributed to differences in incentive value assigned to the A symbol, rather than stimulus-independent differences between hemifields (Chelazzi, et al., 2014).

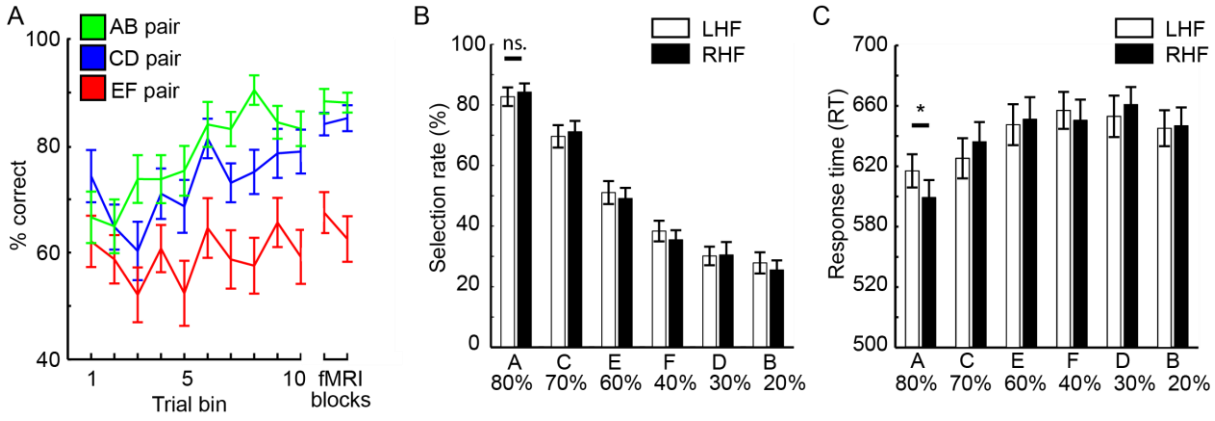

Figure 2. A. Performance as a function of training (Mean±SEM). Performance improved quicker for the AB and CD pairs where the difference in reward probability between the symbols was largest. B. Selection rate during testing with novel pairs (Mean±SEM). On average the A symbol was not selected more frequently in one hemifield over the other. However, symbols associated with higher reward probabilities during training were more frequently selected while symbols associated with lower reward probabilities were more frequently rejected, indicating a generalization of learned values to the novel pairs (see section 3.1.2.2 for statistics). C. Response times (RTs) during testing with novel pairs (Mean±SEM). RTs were fastest for symbols associated with high reward probabilities during training (see section 3.1.2.2 for statistics), and also faster when the A symbol was presented to the RHF (vs. the LHF). LHF/RHF=Left/Right hemifield. *p <0.05.

3.1.3 Modelling results

Figure 3A displays model-derived performance for the different pairs as a function of training. Of note, while the model fits the behavioural data well for AB and CD pairs, it overestimates performance for EF-pairs. This overestimation is caused by fitting one exploration/exploitation parameter β across conditions with different difficulties. Specifically, because performance in the AB and CD pairs is less exploratory than in the more difficult EF pairs, the β parameter will reflect the less exploratory performance found in the majority of the trials (i.e. together AB and CD pairs make up 2/3^{rds} of the trials). This bias results in an underestimation of exploration in EF pairs which causes an overestimation of performance (see Aberg, et al., 2015). Model-derived Q-values at the end of training, for each symbol and each side, are displayed in Figure 3B. The Q-values for the A symbol differed significantly between the hemifields, with a higher Q-value in the RHF as compared to the LHF [$Q_{A\ RVF}$=0.685±0.042 (SEM); $Q_{A\ LVF}$=0.624±0.042 (SEM), p=0.0370]. This result goes well in hand with the finding that RTs for the A symbol were on average faster in the RHF, indicating that the A

symbol had acquired a higher incentive value when presented in the RHF. To determine whether Q-values were generally larger in one hemifield, a repeated measures ANOVA with factors Side (LHF, RHF) and Symbol (A,B,C,D,E,F) was conducted and revealed a significant effect of Symbol [$F(5,352)=153.099$, $p<0.0001$], indicating that symbols with higher reward probabilities had acquired the highest values at the end of training. There was no main effect of Side or interaction between Side and Symbol [both p-values>0.303].

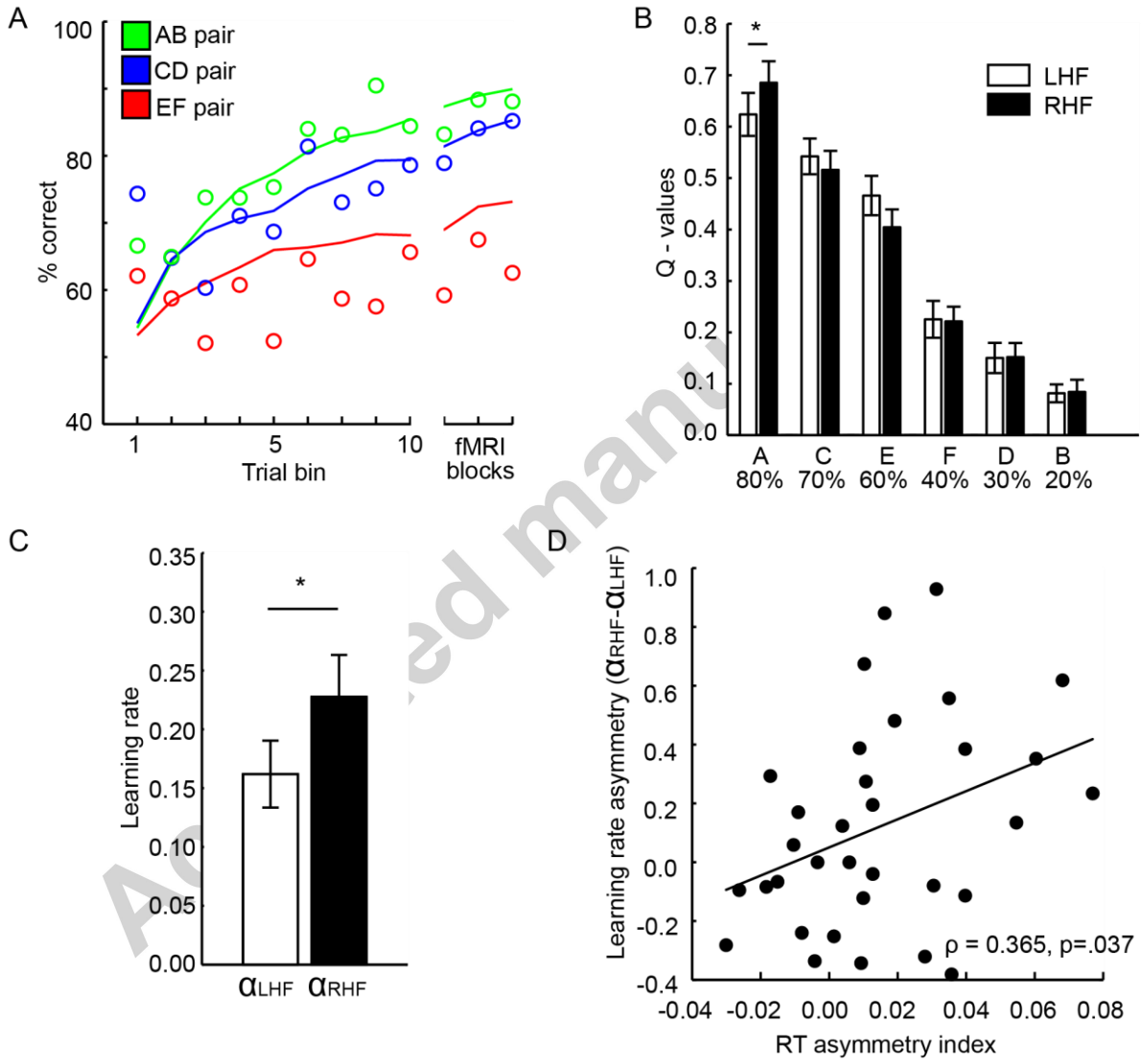

Figure 3. A. Model-derived performance as a function of training. Lines represent the learning curves generated by the model while circles represent actual data. B. Model-derived Q-values at the end of training (Mean±SEM). Q-values for the A symbol in the RHF were significantly larger as compared to the LHF (see section 3.1.3 for statistics). Moreover, symbols associated with higher reward probabilities had acquired higher Q-values. * p<0.05. C. Learning rates for the different

hemifields (Mean±SEM). The learning rates for the RHF were significantly larger than for the LHF. * p<0.05 D. Correlation between post-learning biases and differences in learning rates between hemifields. Participants displaying faster RTs in the RHF also displayed larger learning rates in the RHF. For display purposes, both unranked data and the linear regression slope are shown. LHF/RHF=Left/Right hemifield.

Model-derived learning rates are displayed in Figure 3C. The learning rate was larger for items selected in the RHF as compared to the LHF [α_{RHF}=0.228±0.036 (SEM); α_{HVF}=0.162±0.029 (SEM), p=0.044]. Additionally, hemifield differences in learning rates correlated with the RT asymmetry index [Fig 3D; ρ=0.365, p=0.037] indicating that participants displaying faster RTs in the RHF (vs. the LHF) during testing also displayed larger learning rates in the RHF (vs. the LHF) during training. This latter result provides a computational link between learning biases between hemifields and the expression of post-learning hemifield biases. There was no significant correlation with the selection rate asymmetry index [ρ=-0.118, p=0.488].

3.2 Functional MRI results

We first assessed individual hemispheric asymmetries in DA function by estimating the differences in the neural response to positive and negative feedback during training (Aberg, et al., 2015). Next, we used the computational approach to study the neural correlates of hemifield-specific prediction errors. Finally, we tested for neural activity evoked by hemifield differences in incentive values during the post-learning test phase.

3.2.1 Training

3.2.1.1 Reward processing

Brain regions showing increased response to positive versus negative feedback are displayed in Figure 4A. Activity in the bilateral ventral striatum was higher for positive versus negative feedback

[left ventral striatum: peak voxel MNI x=-12 y=8 z=-11, $t(1,32)=3.966$, $p_{SVC}=0.005$; right ventral striatum peak voxel MNI x=9 y=11 z=-11, $t(1,32)=6.417$, $p_{SVC}<0.001$].

To assess hemispheric differences in reward responses, we estimated an asymmetry index for each individual by calculating the difference in neural response to positive and negative feedback between the right and left ventral striatum (see Section 2.5.2.1). For each participant, the neural response to positive and negative feedback was extracted from two 3 mm radius spheres centred on the peak voxels of the above reported positive>negative feedback contrast in the left and the right ventral striatum. Individual differences in hemispheric reward asymmetry are shown in Figure 4B with 14 participants displaying relatively larger reward responses in the left as compared to the right ventral striatum while 19 participants displayed the reverse pattern.

Next, we tested whether the hemispheric reward asymmetry correlated with the two indexes of hemispatial learning biases. There was no significant correlation with the selection asymmetry index [Fig 4C; $\rho=0.184$, $p=0.323$]. By contrast, there was a significant correlation with the RT asymmetry index [Fig 4D; $\rho=-0.338$, $p=0.046$] because participants with relatively larger reward responses in the left (vs. the right) ventral striatum responded faster when the A symbol was presented to the RHF (vs. the LHF). Given that reward responses in the ventral striatum are closely associated with DA function (Knutson & Gibbs, 2007), this result provides unprecedented evidence linking hemispheric asymmetries in DA function to hemispatial learning biases.

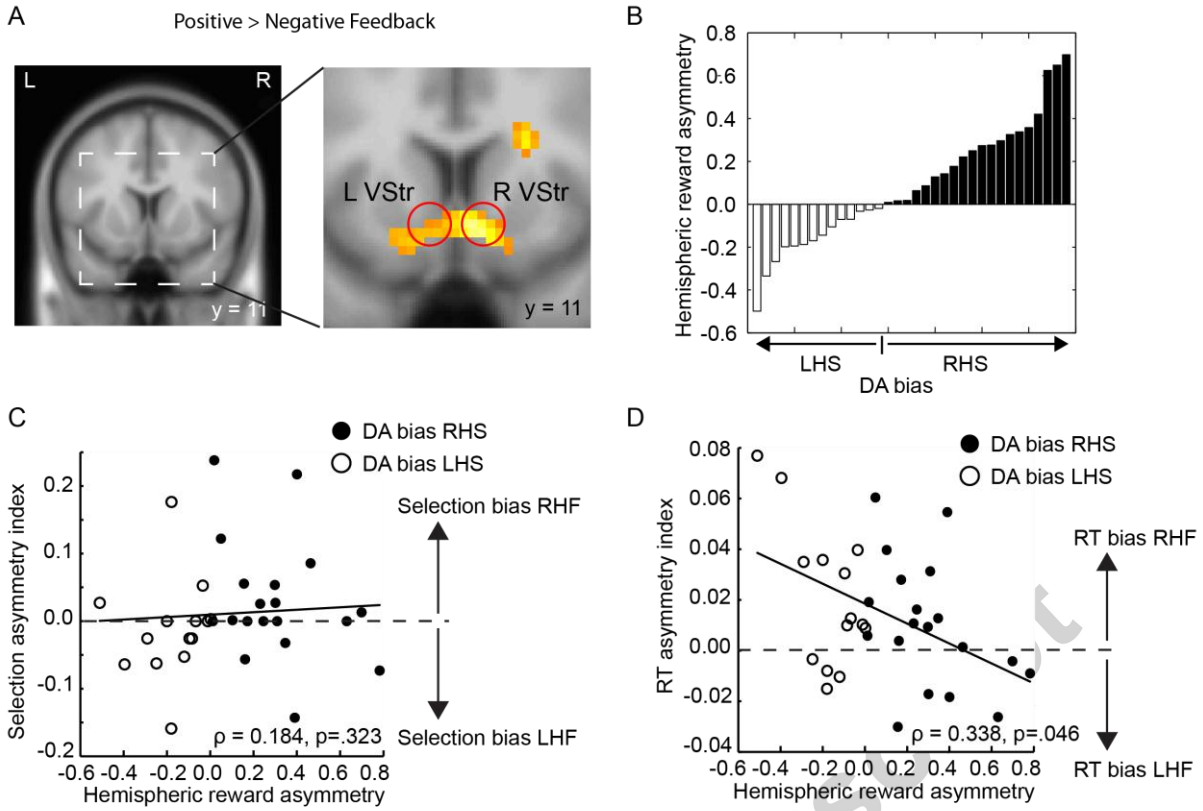

Figure 4. A. Increased neural activity to positive versus negative feedback. The neural response to positive feedback was larger as compared to negative feedback in the bilateral ventral striatum (VStr). Red circles indicate a priori ventral striatum ROIs. For display purposes, activity is displayed at an uncorrected threshold p=0.005. B. The distribution of the hemispheric asymmetry in reward processing. Fourteen participants had increased activity in the left (vs. the right) ventral striatum for positive (vs. negative) feedback while 19 participants displayed the reverse trend. C. Correlation between the selection asymmetry index and the hemispheric reward asymmetry. There was no significant correlation. For display purposes, unranked data are shown as well as the linear regression slope. D. Correlation between the RT asymmetry index and the hemispheric reward asymmetry. Participants displaying faster RTs in the RHF (vs. the LHF) also displayed larger reward responses in the left (vs. the right) ventral striatum. For display purposes, both unranked data and the linear regression slope are shown. LHS/RHS=Left/Right hemisphere. LHF/RHF=Left/Right hemifield.

3.2.1.2 Prediction errors (PEs)

Next, we sought to identify brain regions involved in the neural encoding of hemifield-specific PEs (i.e. PE_{RHF}, PE_{LHF}). Surprisingly, no brain region correlated with PE_{LHF}, even at a threshold of p=0.005. By contrast, PE_{RHF} correlated significantly with activity in the left ventral striatum [Fig 5B; peak voxel MNI x=-6 y=5 z=-8, t(32)=3.728, p_{SVC}=0.009] and the VTA [peak voxel MNI x=6 y=-16 z=-14, t(32)=3.972, p_{SVC}=0.004]. We also tested whether PE_{RHF} was relatively better encoded as compared to

PE_{LHF} (and vice versa). Neither the left nor the right ventral striatum showed any evidence of relatively better encoding of PE_{RHF} or PE_{LHF}, even at a threshold of p=0.005. However, the left ventral striatum showed relatively better encoding of PE_{RHF} as compared to PE_{LHF} with an increasing RT asymmetry index [Fig 5C,D; MNI x=-12 y=11 z=-11, t(32)=3.196, p_{SVC}=0.026]. To pinpoint the exact contribution of the different PEs to this relationship, separate correlations were calculated between the RT asymmetry index and extracted beta parameters for PE_{RHF} and PE_{LHF} for the peak voxel activity within the left ventral striatum (MNI x=-12 y=11 z=-11). While the RT asymmetry index showed a positive correlation with PE_{RHF} [ρ=0.352, p=0.047], it showed a negative correlation with PE_{LHF} [ρ=-0.476, p=0.009]. These results support the notion that hemispatial reward learning is largely determined by reward and reinforcement learning mechanisms in the contralateral striatum.

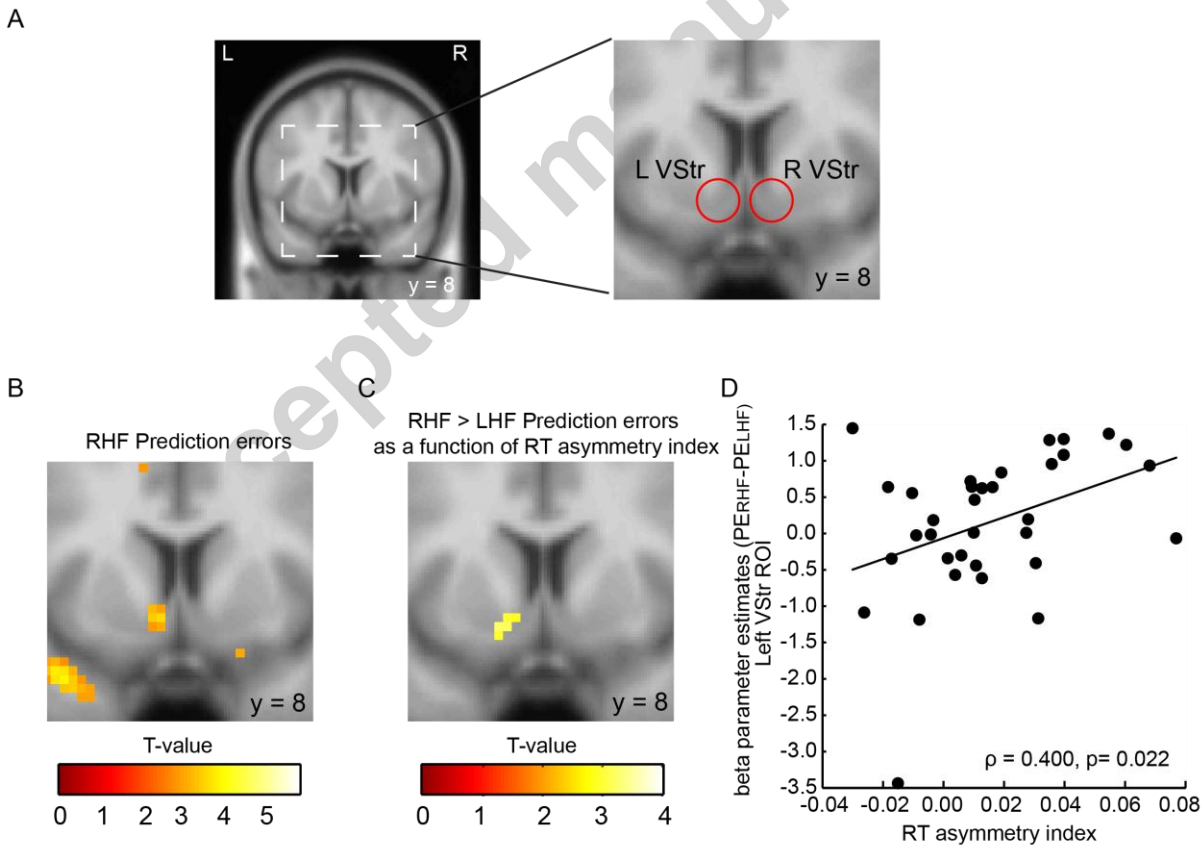

Figure 5. A. Ventral striatum ROIs, denoted by red circles. B. Neural correlates of RHF prediction errors (PE_{RHF}). Activity in the contralateral left ventral striatum (VStr) correlated significantly with PE_{RHF}. Of note, no neural activity correlated with

PEs in the LHF (PE_{LVF}). For display purposes, activity displayed at an uncorrected threshold p=0.005. C. Neural correlates of the relative encoding of PE_{RHF} and PE_{LVF} as a function of RT asymmetry index. Activity in the left ventral striatum tracked the relative encoding of PEs in the different visual hemifields as a function of RT asymmetry index. For display purposes, activity is displayed at an uncorrected threshold p=0.005. D. The average activity in the left ventral striatum ROI tracked the relative encoding of PE_{RHF} and PE_{LVF} as a function of RT asymmetry index. Participants with faster RTs in the RHF (vs. the LHF) showed better encoding of PEs in the RHF (vs. the LHF). For display purposes, both unranked data and the linear regression slope are shown. LHF/RHF=Left/Right hemifield.

3.2.2 Testing

3.2.2.1 Learned values

Next, we tested whether brain regions known to mediate motivation and incentive value contributed to the expression of a hemifield bias during testing. Presenting the A symbol to the RHF vs. the LHF (A_{RHF}-A_{LHF}) increased activity in the VTA [Fig 6B; MNI x=0 y=-13 z=-14, t(32)=3.437, p_{SVC}=0.009; MNI x=3 y=-16 z=-17, t(32)=3.320, p_{SVC}=0.012]. As the VTA has been strongly implicated in mediating incentive motivation and decreased RTs (Adcock, et al., 2006; Arsenault, Rima, Stemmann, & Vanduffel, 2014; Knutson, et al., 2005; Yun, Wakabayashi, Fields, & Nicola, 2004), this result goes well in hand with the observed overall decrease in RTs when the A symbol was presented to the RHF (vs. LHF). Moreover, activity in the VTA also increased when the model-derived Q-value was increasingly larger for the symbol presented to the RHF (Q_{RHF}-Q_{LHF})[Fig 6C; MNI x=3 y=-19 z=-11, t(32)=2.881, p_{SVC}= 0.028], indicating that the VTA is not only sensitive to hemifield differences in reward value between pairs, but also tracks subtle differences in relative value within pairs of symbols. These results thus suggest that the expression of post-learning hemispatial biases is influenced by activity in dopaminergic networks commonly believed to underpin incentive motivation. Of note, no evidence indicated that activity in the ventral striatum was modulated by hemifield-differences in symbol values.

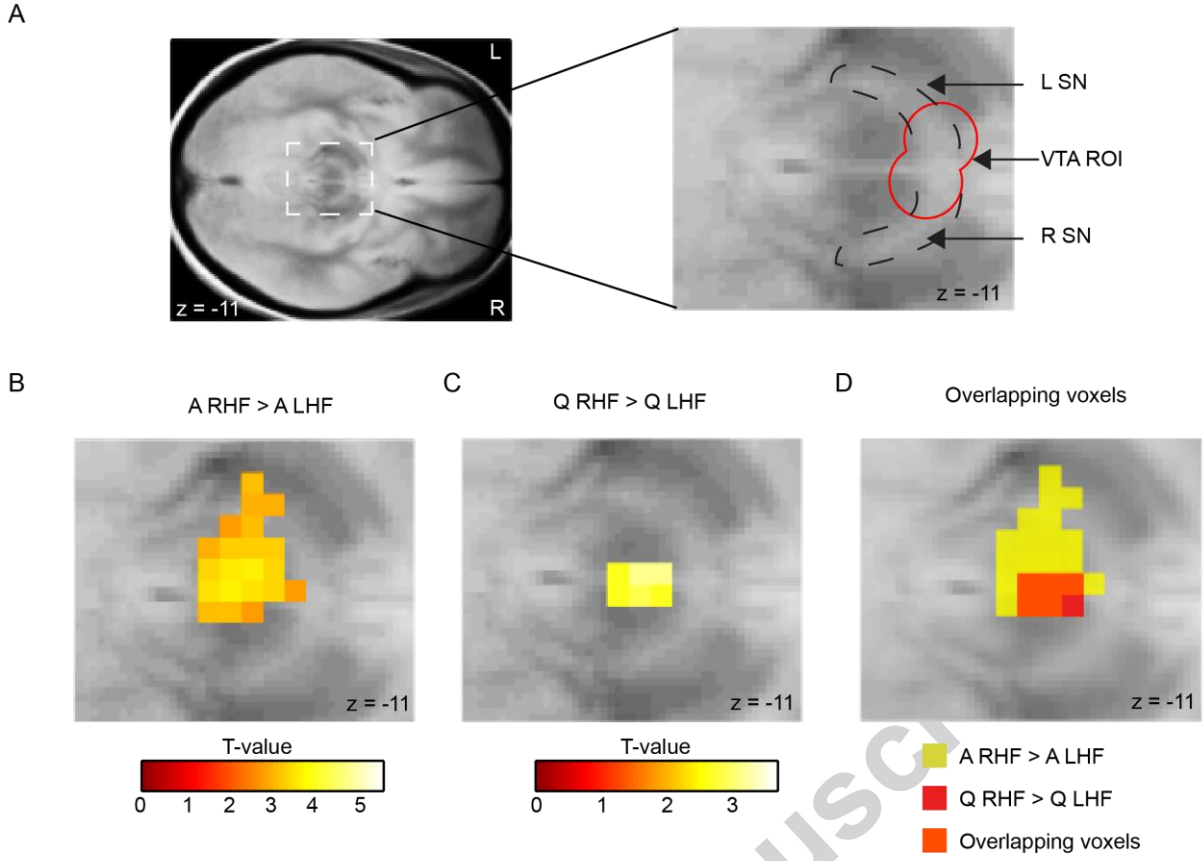

Figure 6. A. VTA ROI, delimited by a red line and overlaid on a proton density (PD) structural scan averaged across participants. The VTA is located medially to the substantia nigra (SN) which can be identified as a white strip in PD images. B. Neural activity as a function of the difference between A_{RHF} and A_{LHF} between pairs during testing. Activity increased in the VTA when the A symbol was presented to the RHF as compared to when it was presented to the LHF. For display purposes, activity displayed at an uncorrected threshold p=0.005. C. Neural activity as a function of the difference between Q_{RHF} and Q_{LHF} within pairs. Activity in the VTA increased when the relative difference in Q-value between the RHF and the LHF symbol increased. For display purposes, activity displayed at an uncorrected threshold p=0.005. D. Overlapping voxels. Some voxels in the VTA increased activity both when the A symbol was presented to the RHF (vs. the LHF) and when the Q-value of the symbol in the RHF increased relative the symbol in the LHF. LHF/RHF=Left/Right hemifield.

4. Discussion

In the present study, we tested whether the expression of reward-related spatial biases implicates hemispheric asymmetries in DA function and in underlying computational learning mechanisms. Specifically, we hypothesized that spatial biases in learning the reward value of symbols presented to different hemifields would relate to hemispheric asymmetries in reward responses and to the

efficiency of PE signalling. To this end, participants learned reward probabilities of laterally presented symbols in a trial-and-error fashion while being scanned. A computational approach was used to unveil computational and neural biases during learning which may relate to the behavioural expression of hemispatial biases after the learning. The results are discussed in detail below.

4.1 Asymmetries in striatal reward responses predicts spatial learning biases

Hemispheric asymmetries in DA function have previously been linked to orienting biases because animals and humans tend to orient away from the hemisphere with more DA activity (Mohr, et al., 2004; Nash, et al., 2010; Tomer, et al., 2013). Yet, while one experiment in rats showed that individual orienting biases in a T-maze were associated with deficits when learning to enter the non-preferred arm in the maze (Andrade, et al., 2001), it remains unclear whether hemispheric asymmetries in DA function contributes to hemispatial learning biases. The present study reports a relationship between hemispheric asymmetries in reward processing in the ventral striatum and biases in learning reward values of symbols presented to different hemifields in humans. Because striatal implication to reward processing relies on DA function (Knutson & Gibbs, 2007), the present results provide a first indication that biases in hemispatial reward learning may be determined by hemispheric asymmetries in striatal DA function (see also section 4.2).

Importantly, biases in response vigour (RT asymmetry index) rather than symbol selection (selection asymmetry index) correlated with asymmetries in hemispheric reward responses. These results fit well with the characteristics of the task in which reward probabilities were presented as fixed, discrete values. Indeed, although asymmetries in hemispatial learning could induce distinct values for the same symbol presented to different hemifields, this may not translate into different selection rates as the intra-symbol value difference (i.e. A_{LHF} vs. A_{RHF}) may be smaller than inter-symbol value differences (i.e. A_{LHF} vs. C_{LHF} and A_{RHF} vs. C_{RHF}), resulting in largely the same selection rates for the same symbol presented to different hemifields. By contrast, RTs represent a continuous

measure which may provide a more sensitive measure of learning-induced differences in incentive values, that may also be expressed implicitly (Schmidt, et al., 2010). Further support for this interpretation comes from the computational and neuroimaging data (see section 4.2 and 4.3), which indicate that the RT asymmetry index, and not the selection rate asymmetry index, relates to biases in reinforcement learning mechanisms.

4.2 Biased computational reinforcement learning relates to spatial learning biases

To better understand spatial learning biases, we used a computational approach to model the pattern of behavioural responses during the learning of the probabilistic selection task, separately for symbols presented to the RVF and LVF. The model revealed that the frequently rewarded A symbol had acquired a significantly higher Q-value in the RHF as compared to the LHF, a result which dovetails with the finding that RTs were generally shorter for the A symbol in the RHF. This finding also provides insights into possible mechanisms contributing to the expression of post-learning spatial biases, as discussed next.

Previous research reported that RTs for cues associated with high rewards are typically faster as compared to RTs for cues associated with low rewards (Adcock, et al., 2006; Knutson, Fong, et al., 2001; Mir, et al., 2011), suggesting that large Q-values (i.e. reflecting large expected rewards) are linked to fast RTs through the concept of incentive motivation. Fluctuations in incentive motivation were also found to correlate with activity in brain regions involved in attention and motor function, as well as with activity in the dopaminergic midbrain and its projection sites such as the ventral striatum (Knutson, Adams, Fong, & Hommer, 2001; Knutson, et al., 2005; Tobler, et al., 2005). These findings, together with findings highlighting a crucial role for DA in activational aspects of behaviour (Robbins & Everitt, 1992, 2007), have led researchers to suggest that the impact of motivation on behaviour is mediated by DA (Chakravarthy, Joseph, & Bapi, 2010; Salamone & Correa, 2012; Wise, 2004). Several studies have also shown that the same neural circuitry may also represent the

expected value of stimuli presented to one hemifield, but largely restricted to the contralateral hemisphere (Gershman, Pesaran, & Daw, 2009; Palminteri, et al., 2009; Wunderlich, Rangel, & O'Doherty, 2009), suggesting the possibility of unilaterally activating the brain's motivational system. This notion was recently confirmed in a study where unilateral deep brain stimulation of the subthalamic nucleus, a procedure mimicking the impact of DA enhancers, boosted motivational processes within the stimulated hemisphere, as indicated by greater exerted force for the hand controlled by the stimulated hemisphere following increased monetary incentives (Palminteri, et al., 2013). Moreover, unilateral incentive motivation could also be induced by presenting subliminal rewards to one hemifield, as indicated by increased response vigour (i.e. greater exerted force) for the hand controlled by the hemisphere contralateral (vs. ipsilateral) to the rewarded visual stimulation (Schmidt, et al., 2010). Accordingly, the incentive value of stimuli presented to one hemifield would determine the degree of activation of motivational brain circuitry in the contralateral hemisphere. A derivate of these findings is that differences in incentive values for stimuli presented to different hemifields should cause hemispheric asymmetries in the activation of motivational brain circuits. This suggestion is well in line with the present results, because asymmetric differences in Q-values for the A symbol presented to different hemifields related to asymmetric differences in RTs.

Another interpretation for these results can be offered by the theory of spatial priority (or saliency) maps, which suggests that various types of spatial information are integrated into spatial maps, which can be accessed by attention and motor processes, and thus guide quick orientation towards highly prioritized spatial locations (Chelazzi, et al., 2014; Itti & Koch, 2001). These maps combine information related to the spatial location itself (e.g. attentional cues, previous reward history), as well as information specific to presented stimuli (e.g. emotional saliency, expected reward value). The formation and updating of these maps likely involves reinforcement learning processes (Chelazzi, Perlato, Santandrea, & Della Libera, 2013). Critically, this theory predicts that biases in orienting responses to two separate spatial locations are related to distortions in these

priority maps caused by, for example, differences in expected reward value for stimuli presented at those locations (Chelazzi, et al., 2014). In line with this prediction, we found shorter response times when the highly rewarded A-symbol was presented at the location where it acquired a higher Q-value (here, in the RHF), thus also confirming a contribution of expected reward to spatial priority maps (Navalpakkam, Koch, & Perona, 2009). Because spatial priority maps incorporate aspects of incentive motivation, such as the expected value of stimuli, the two aforementioned interpretations are not mutually exclusive.

The learning rate, which controls how quickly symbol values are updated, was on average higher for the RHF as compared to the LHF. Additionally, participants displaying higher learning rates in the RHF (vs. the LHF) also displayed faster RTs for A symbols in the RHF (vs. the LHF), a result which provides a computational link between hemispatial learning biases and the expression of hemispatial biases after learning. Elevated learning rates are particularly beneficial in highly deterministic contexts when outcomes frequently represent actual values. In the present study, the reward probability of A symbol (0.8) is most closely linked to the reward outcomes, suggesting that it should be most sensitive to differences in learning rates between hemifields. Corroborating this notion, the data indicate both higher learning rates and Q-values for the A symbol in the RHF (vs. the LVF). (Note that the B symbol with a reward probability of 0.2 is equally closely coupled to the reward outcome, but the estimation of its Q-value may suffer from a more variable reward history as learning causes participants to select the B symbol less frequently in AB pairs).

Finally, we would like to specify why we investigated individual differences in learning using a computational approach rather than standard behavioural measures, such as mean RT or selections. First of all, it is not clear whether the latter measurements during the training phase would be mainly guided by differences in acquired reward probabilities (i.e. stimulus value), or by other factors such as habitual responding caused by presenting only three pairs of symbols. Additionally, performance could be determined either by selecting the symbol with the highest reward probability (i.e. A in an

AB pair), or by avoiding the symbol with the lowest reward probability (i.e. B in an AB pair). This information cannot be inferred from mean RT or selection measures during the training phase (but can be inferred from the test phase where behaviour could only be guided by differences in the reward probability, i.e. expected value assigned to each symbol) (Frank, et al., 2004). By contrast, the computational approach estimated one value for each symbol based solely on its previous reward history, making it insensitive to other confounding factors, such as those mentioned above. It also provided one estimate of the Q-value acquired for each symbol during the whole training, a summary value which we showed here correlated with individual RT biases in the subsequent test phase. This finding further evidence that the computational model successfully extracted relevant information from one modality and session (i.e. selection during training) to predict performance for another modality and session (i.e. RT during testing). Therefore, the computational approach used in the present experimental context offered a more refined, accurate, and meaningful way of capturing essential determinants of individual behaviour expressed during the training phase, as compared to standard behavioural measures. Moreover, when combined with neuroimaging methods, this approach can also provide a detailed description of the development of behavioural biases both at the mechanistic and at neural levels, i.e. when combined with neuroimaging methods.

4.3 Biased neural reinforcement learning relates to spatial learning biases

To investigate the neural correlates of hemispatial learning and the development of hemispatial biases, we analysed the fMRI data using separate PEs for symbols selected in the RHF and the LHF (i.e. PE_{RHF} and PE_{LHF}), as provided by the computational model. Activity in the VTA and the left ventral striatum correlated with PE_{RHF}. Additionally, the relative encoding of PE_{RHF} (vs. PE_{LHF}) in the left ventral striatum correlated with the hemispatial bias expressed after learning.

A stronger correlation between PEs and brain activity indicates a better neural representation of the learning signal (Schonberg, Daw, Joel, & O'Doherty, 2007), which is causally linked to learning efficacy (Aquili, 2014; Schonberg, et al., 2007; Steinberg, et al., 2013). These results therefore

support, at least partially, the notion that the striatum contributes to hemispatial learning by encoding the learning signal needed for behavioural adaptation in the contralateral hemifield. Additionally, they provide evidence that a relatively better neural representation of the learning signal related to one behaviour leads to behavioural biases favouring that behaviour over other behaviours (see also Aberg, et al., 2015).

Hemisphere-specific reward learning mechanisms, as substantiated here, may also explain some recent reports of deficits in hemispatial learning and hemispatial reward processing, specifically in patients with unilateral lesions extending into the striatum (Lucas, et al., 2013; Malhotra, et al., 2013), i.e. such learning deficits could have been caused by disruptions of striatal reward- and reinforcement learning processes. Moreover, the degree of rightward-directed hemispatial learning bias was associated with both increased reward processing in the ventral striatum, known to be modulated by DA (Knutson & Gibbs, 2007; van der Vegt, et al., 2013), and relatively better neural encoding of PEs in the contralateral left ventral striatum, which is consistent with pharmacological studies showing that the neural encoding of PEs in the ventral striatum is influenced by the level of DA function (Chowdhury, et al., 2013; Jocham, et al., 2011; Pessiglione, et al., 2006). This reasoning may also provide an alternative explanation for the results of a recent study by Palminteri, et al., (2013), in which unilateral increases in DA function, achieved through deep brain stimulation of the subthalamic nucleus (DBS-STN), enhanced reward learning for stimuli presented in the hemifield contralateral (vs. ipsilateral) to the stimulated hemisphere. Using a computational approach, the authors showed that spatial learning biases may have been mediated by increased reward processing in the DBS-STN stimulated hemisphere. Based on the present results, we may speculate that better reward learning in one hemifield reported by Palminteri, et al., (2013) was caused by enhanced PE encoding in the ventral striatum of the contralateral, DBS-STN-stimulated hemisphere.

Another interpretation of the links between spatial learning biases and hemisphere-specific reinforcement learning mechanisms, as we report here, can be derived from research on multi-

effector action reinforcement which posits that different actions are learned and associated to the effectors performing them (Daw, 2014; Madlon-Kay, Pesaran, & Daw, 2013; Wunderlich, et al., 2009). For example, PEs for actions performed by the right or left hand are predominantly represented in the contralateral ventral striatum (Gershman, et al., 2009; Madlon-Kay, et al., 2013; Palminteri, et al., 2009). Because all effectors used to respond in the present study were controlled by the left hemisphere (i.e. the fingers of the right hand), it could be suggested that PEs should be represented predominantly in the left ventral striatum. Consistent with this suggestion, the left ventral striatum displayed a positive correlation between the encoding of PEs related to the right hemifield and the expression of post-learning spatial biases (i.e. the RT asymmetry index), while it showed a negative correlation between the RT asymmetry index and the encoding of PEs related to the left hemifield. No evidence indicated the involvement of the right ventral striatum in encoding PEs. Moreover, this result extends previous literature by showing how asymmetric encoding of PEs influences subsequent behaviours. Indeed, although some studies found that PEs related to using the left and right hand were predominantly represented in the right and left ventral striatum, respectively, any related behavioural consequence has remained unknown (Gershman, et al., 2009; Madlon-Kay, et al., 2013). By contrast, here we show that encoding of PEs related to one hemifield (vs. the other) was associated with the subsequent expression of spatial biases. This result also relates to the recent observation that better neural encoding of PEs related to approach (vs. avoidance) learning promoted approach (vs. avoidance) behaviours in a subsequent test phase (Aberg, et al., 2015).

Orienting biases (Andrade, et al., 2001; Tomer, 2008), hemispheric asymmetries in reward processing (Aberg, et al., 2015), and DA function (Tomer, Goldstein, Wang, Wong, & Volkow, 2008) have been related to behavioural predispositions and traits. Our results show that hemispheric asymmetries in reward processing, pertaining to asymmetries in DA function, implicated biases in spatial learning, both at a neural and mechanistic level. While these results lend further evidence to the notion that the expression of trait-like behavioural biases may be maintained and reinforced

through trait-related modulations of learning processes (Indovina, et al., 2011; Olsson, et al., 2013; Sharot, et al., 2011), it also raises the question concerning cause and effect. Some evidence indicate that orienting biases are linked to the expression of dopaminergic genes (Greene, Robertson, Gill, & Bellgrove, 2010; Newman, O'Connell, Nathan, & Bellgrove, 2012; Zozulinsky, et al., 2014). Additionally, there may also be a genetic basis to hemispheric asymmetries in DA function (Pohjalainen, et al., 1998; see Zozulinsky, et al., 2014 for a discussion). One plausible explanation would therefore be that genetic predispositions cause asymmetric DA function which would in turn determine the expression of orienting biases. A similar argument can be made for the expression of spatial learning biases. Accordingly, both orienting biases and spatial learning biases may be underpinned by genetically determined hemispheric asymmetries in DA function. However, the present study highlights an important, additional factor, namely that asymmetric DA function also contributes to the maintenance of spatial biases by modulating the spatial learning of new information.

4.4 The VTA mediates motivational biases between hemifields

During testing with novel pairs, activity in the VTA increased when the Q-value was larger for the symbol presented to the RHF (vs. the LHF) as well as when the A symbol was presented to the RHF (vs. the LHF). The VTA has been associated with representing incentive values and incentive motivation (Arsenault, et al., 2014; Knutson, et al., 2005; Tobler, et al., 2005). For example, neural activity in the VTA increases for cues associated with high (vs. low) rewards and this increase is also related to decreased RTs (Adcock, et al., 2006; Knutson, Fong, et al., 2001). Increased VTA activity when the A symbol is presented to the RHF (vs. the LHF) therefore goes well in hand with the finding that the A symbol had acquired higher Q-values and increased incentive motivation, as reflected in faster RTs, when presented to the RHF (vs. the LHF).

While incentive motivation can be increased in one hemisphere separately (Palminteri, et al., 2013; Schmidt, et al., 2010), and the neural representation of reward values for lateralized presented

objects are located in the contralateral hemisphere (Gershman, et al., 2009; Palminteri, et al., 2009; Wunderlich, et al., 2009), it remains unclear whether the neural correlates of hemifield-specific incentive motivation are also lateralized. The present results demonstrate that subtle differences in incentive motivation between hemifields, implicating lateralized brain structures during learning (i.e. the ventral striatum), and expressed by hemifield differences in incentive value and response vigour during testing, are mediated by brain circuits in the midbrain (i.e. the VTA).

4.5 Limitations

We acknowledge a potential limitation of the present study. DA function here was estimated by the strength of BOLD responses during reward processing. The BOLD response is not a direct measure of DA release or DA neuron activity, and does not discriminate between neural activity evoked by other neuromodulators and neurotransmitters, such as acetylcholine, GABA, or glutamate, which may also influence activity in the striatum and the midbrain and have been associated with modulations of spatial attention (Knudsen, 2011; Stormer, Passow, Biesenack, & Li, 2012). However, transient modulations of fMRI signal across the mesolimbic DA system during reward processing has previously been demonstrated to be highly consistent with neural recordings in animals (D'Ardenne, et al., 2008; Knutson, Fong, et al., 2001; Schultz, Dayan, & Montague, 1997; Tobler, et al., 2005). Additionally, in line with studies using more direct measures of DA function, such as PET (Tomer, et al., 2014) and DA neuron loss in Parkinson's disease patients (Maril, Hassin-Baer, Cohen, & Tomer, 2013), our recent fMRI data evidence a link between hemispheric differences in DA function and biases in approach and avoidance learning (Aberg, et al., 2015). Note that, for ethical reasons, more invasive procedures, such as PET, are not easily available for studying healthy controls in many countries, making this fMRI-based approach an attractive alternative for estimating individual differences in DA function.

4.6 Conclusion

The present study provides computational and brain imaging evidence linking functional hemispheric asymmetries in the reward system to spatial learning biases in humans. In particular, our results suggest that pre-existing individual neuronal reward-processing biases influence hemisphere-specific reinforcement learning mechanisms. Beyond hemispatial learning, the present results may also have clinical implications. Specifically, some individuals are at higher risk of developing mental disorders (Chambers, Power, & Durham, 2004; Indovina, et al., 2011), and recent evidence links the development and maintenance of mental disorders to biased learning processes (Itzhak, Perez-Lanza, & Liddie, 2014; Mineka & Oehlberg, 2008). Therefore, studying individual differences in the neural mechanisms underlying learning may be relevant for understanding why some individuals may be more likely than others to develop mental disorders.

Acknowledgements

This work was supported by the Swiss National Science Foundation (grant numbers: 320030-159862 and 320030-135653) and by the Swiss Center for Affective Sciences.

References

Aberg, K., Doell, C., & Schwartz, S. (2015). Hemispheric asymmetries in striatal reward responses relate to approach-avoidance learning and encoding of positive-negative prediction errors in dopaminergic midbrain regions. The Journal of Neuroscience, 35, 14491-14500.

Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. Neuroimage, 31, 790-795.

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron, 50, 507-517.

Andrade, C., Alwarshetty, M., Sudha, S., & Suresh Chandra, J. (2001). Effect of innate direction bias on T-maze learning in rats: implications for research. J Neurosci Methods, 110, 31-35.

Aquili, L. (2014). The causal role between phasic midbrain dopamine signals and learning. Front Behav Neurosci, 8, 139.

Arsenault, J. T., Rima, S., Stemmann, H., & Vanduffel, W. (2014). Role of the primate ventral tegmental area in reinforcement and motivation. Curr Biol, 24, 1347-1353.

Barone, P., Bankiewicz, K. S., Corsini, G. U., Kopin, I. J., & Chase, T. N. (1987). Dopaminergic mechanisms in hemiparkinsonian monkeys. Neurology, 37, 1592-1595.

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

Bracha, H. S., Shults, C., Glick, S. D., & Kleinman, J. E. (1987). Spontaneous asymmetric circling behavior in hemi-parkinsonism; a human equivalent of the lesioned-circling rodent behavior. Life Sci, 40, 1127-1130.

Cha, J., Carlson, J. M., Dedora, D. J., Greenberg, T., Proudfit, G. H., & Mujica-Parodi, L. R. (2014). Hyper-reactive human ventral tegmental area and aberrant mesocorticolimbic connectivity in overgeneralization of fear in generalized anxiety disorder. J Neurosci, 34, 5855-5860.

Chakravarthy, V. S., Joseph, D., & Bapi, R. S. (2010). What do the basal ganglia do? A modeling perspective. Biol Cybern, 103, 237-253.

Chambers, J. A., Power, K. G., & Durham, R. C. (2004). The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of Generalized Anxiety Disorder. J Anxiety Disord, 18, 587-607.

Chelazzi, L., Estocinova, J., Calletti, R., Lo Gerfo, E., Sani, I., Della Libera, C., & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. J Neurosci, 34, 8594-8604.

Chelazzi, L., Perlato, A., Santandrea, E., & Della Libera, C. (2013). Rewards teach visual selective attention. Vision Res, 85, 58-72.

Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Duzel, E., & Dolan, R. J. (2013). Dopamine restores reward prediction errors in old age. Nat Neurosci, 16, 648-653.

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

Daw, N. D. (2014). Advanced Reinforcement Learning. In P. W. Glimcher & E. Fehr (Eds.), Neuroeconomics (Second Edition ed.). New York: Academic Press.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci U S A, 104, 16311-16316.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science, 306, 1940-1943.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. Neuron, 47, 495-501.

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

Glascher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron, 66, 585-595.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc Natl Acad Sci U S A, 108 Suppl 3, 15647-15654.

Greene, C. M., Robertson, I. H., Gill, M., & Bellgrove, M. A. (2010). Dopaminergic genotype influences spatial bias in healthy adults. Neuropsychologia, 48, 2458-2464.

Howell, D. (2013). Statistical Methods for Psychology (8th ed.). Belmont, CA: Wadsworth.

Indovina, I., Robbins, T. W., Nunez-Elizalde, A. O., Dunn, B. D., & Bishop, S. J. (2011). Fear-conditioning mechanisms associated with trait vulnerability to anxiety in humans. Neuron, 69, 563-571.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. Nat Rev Neurosci, 2, 194-203.

Itzhak, Y., Perez-Lanza, D., & Liddie, S. (2014). The strength of aversive and appetitive associations and maladaptive behaviors. IUBMB Life, 66, 559-571.

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. J Neurosci, 31, 1606-1613.

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2009). The strength of anticipatory spatial biasing predicts target discrimination at attended locations: a high-density EEG study. Eur J Neurosci, 30, 2224-2234.

Kinsbourne, M. (1970). The cerebral basis of lateral asymmetries in attention. Acta Psychol (Amst), 33, 193-201.

Knudsen, E. I. (2011). Control from below: the role of a midbrain network in spatial attention. Eur J Neurosci, 33, 1961-1972.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J Neurosci, 21, RC159.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport, 12, 3683-3687.

Knutson, B., & Gibbs, S. E. B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. Psychopharmacology (Berl), 191, 813-822.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. J Neurosci, 25, 4806-4812.

Lee, A. C., Harris, J. P., Atkinson, E. A., & Fowler, M. S. (2001). Evidence from a line bisection task for visuospatial neglect in left hemiparkinson's disease. Vision Res, 41, 2677-2686.

Lucas, N., Schwartz, S., Leroy, R., Pavin, S., Diserens, K., & Vuilleumier, P. (2013). Gambling against neglect: unconscious spatial biases induced by reward reinforcement in healthy people and brain-damaged patients. Cortex, 49, 2616-2627.

Madlon-Kay, S., Pesaran, B., & Daw, N. D. (2013). Action selection in multi-effector decision making. Neuroimage, 70, 66-79.

Malhotra, P. A., Soto, D., Li, K., & Russell, C. (2013). Reward modulates spatial neglect. J Neurol Neurosurg Psychiatry, 84, 366-369.

Maril, S., Hassin-Baer, S., Cohen, O., & Tomer, R. (2013). Effects of asymmetric dopamine depletion on sensitivity to rewarding and aversive stimuli in Parkinson's disease. Neuropsychologica, 818-824.

Mineka, S., & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. Acta Psychol (Amst), 127, 567-580.

Mir, P., Trender-Gerhard, I., Edwards, M. J., Schneider, S. A., Bhatia, K. P., & Jahanshahi, M. (2011). Motivation and movement: the effect of monetary incentive on performance speed. Exp Brain Res, 209, 551-559.

Mohr, C., Brugger, P., Bracha, H. S., Landis, T., & Viaud-Delmon, I. (2004). Human side preferences in three different whole-body movement tasks. Behav Brain Res, 151, 321-326.

Nash, K., McGregor, I., & Inzlicht, M. (2010). Line bisection as a neural marker of approach motivation. Psychophysiology, 47, 979-983.

Navalpakkam, V., Koch, C., & Perona, P. (2009). Homo economicus in visual search. J Vis, 9, 31 31-16.

Neto, L. L., Oliveira, E., Correia, F., & Ferreira, A. G. (2008). The human nucleus accumbens: where is it? A stereotactic, anatomical and magnetic resonance imaging study. Neuromodulation, 11, 13-22.

Newman, D. P., O'Connell, R. G., & Bellgrove, M. A. (2013). Linking time-on-task, spatial bias and hemispheric activation asymmetry: a neural correlate of rightward attention drift. Neuropsychologia, 51, 1215-1223.

Newman, D. P., O'Connell, R. G., Nathan, P. J., & Bellgrove, M. A. (2012). Dopamine transporter genotype predicts attentional asymmetry in healthy adults. Neuropsychologia, 50, 2823-2829.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl), 191, 507-520.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. Neuron, 38, 329-337.

Olsson, A., Carmona, S., Downey, G., Bolger, N., & Ochsner, K. N. (2013). Learning biases underlying individual differences in sensitivity to social rejection. Emotion, 13, 616-621.

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., & Pessiglione, M. (2009). Brain hemispheres selectively track the expected value of contralateral options. J Neurosci, 29, 13465-13472.

Palminteri, S., Serra, G., Buot, A., Schmidt, L., Welter, M. L., & Pessiglione, M. (2013). Hemispheric dissociation of reward processing in humans: insights from deep brain stimulation. Cortex, 49, 2834-2844.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature, 442, 1042-1045.

Pohjalainen, T., Rinne, J. O., Nagren, K., Lehikoinen, P., Anttila, K., Syvalahti, E. K., & Hietala, J. (1998). The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Mol Psychiatry, 3, 256-260.

Robbins, T. W., & Everitt, B. J. (1992). Functions of dopamine in the dorsal and ventral striatum. Seminars in Neuroscience, 4, 119-127.

Robbins, T. W., & Everitt, B. J. (2007). A role for mesencephalic dopamine in activation: commentary on Berridge (2006). Psychopharmacology (Berl), 191, 433-437.

Salamone, J. D., & Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. Neuron, 76, 470-485.

Schmidt, L., Palminteri, S., Lafargue, G., & Pessiglione, M. (2010). Splitting motivation: unilateral effects of subliminal incentives. Psychol Sci, 21, 977-983.

Schonberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. J Neurosci, 27, 12860-12867.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. Science, 275, 1593-1599.

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. Annu Rev Neurosci, 23, 473-500.

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. Nat Neurosci, 14, 1475-1479.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci, 16, 966-973.

Stormer, V. S., Passow, S., Biesenack, J., & Li, S. C. (2012). Dopaminergic and cholinergic modulations of visual-spatial attention and working memory: insights from molecular genetic research and implications for adult cognitive development. Dev Psychol, 48, 875-889.

Sutton, R., & Barto, A. (1998). Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press.

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. J Neurosci, 26, 9494-9502.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. Science, 307, 1642-1645.

Tomer, R. (2008). Attentional bias as trait: correlations with novelty seeking. Neuropsychologia, 46, 2064-2070.

Tomer, R., Goldstein, R. Z., Wang, G. J., Wong, C., & Volkow, N. D. (2008). Incentive motivation is associated with striatal dopamine asymmetry. Biol Psychol, 77, 98-101.

Tomer, R., Slagter, H. A., Christian, B., Fox, A. S., King, C. R., Murali, D., & Davidson, R. J. (2013). Dopamine asymmetries predict orienting bias in healthy individuals. Cereb Cortex, 23, 2899-2904.

Tomer, R., Slagter, H. A., Christian, B. T., Fox, A. S., King, C. R., Murali, D., Gluck, M. A., & Davidson, R. J. (2014). Love to win or hate to lose? Asymmetry of dopamine D2 receptor binding predicts sensitivity to reward vs. punishment. Journal of Cognitive Neuroscience, 1039-1048.

van der Vegt, J., Hulme, O., Zittel, S., Madsen, K., Weiss, M., Buhmann, C., Bloem, B., Münchau, A., & Siebner, H. (2013). Attenuated neural response to gamble outcomes in drug-naive patients with Parkinson's disease. Brain, 136, 1192-1203.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. Perspectives on Psychological Science, 4, 274-290.

Watkins, C., & Dayan, P. (1992). Q-learning. Machine Learning, 279-292.
Wise, R. A. (2004). Dopamine, learning and motivation. Nat Rev Neurosci, 5, 483-494.
Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. Proc Natl Acad Sci U S A, 106, 17199-17204.
Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues. J Neurosci, 24, 2923-2933.
Zozulinsky, P., Greenbaum, L., Brande-Eilat, N., Braun, Y., Shalev, I., & Tomer, R. (2014). Dopamine system genes are associated with orienting bias among healthy individuals. Neuropsychologia, 62, 48-54.

Highlights

- Dopamine (DA) function determine reward and reinforcement learning processes
- Hemispheric asymmetries in DA function determine spatial orienting biases
- We show that asymmetries in DA function also bias hemispatial reward learning
- Model-based fMRI revealed the mechanisms mediating such biases
- Hemispheric asymmetries in striatal function were particularly implicated