Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning

Stefano Palminteri,[1,2] Damian Justo,[2] Céline Jauffret,[2] Beth Pavlicek,[3] Aurélie Dauta,[4] Christine Delmaire,[7] Virginie Czernecki,[5] Carine Karachi,[4] Laurent Capelle,[4] Alexandra Durr,[2,6] and Mathias Pessiglione[1,2,*]

[1]Motivation, Brain and Behavior (MBB) Team, Institut du Cerveau et de la Moelle Épinière (ICM), Hôpital de la Pitié-Salpêtrière, 75013 Paris, France
[2]INSERM UMR 975, CNRS UMR 7225, Université Pierre et Marie Curie (UMPC – Paris 6), 75013 Paris, France
[3]Decision Neuroscience Group, INSEAD, 77300 Fontainebleau, France
[4]Service de Neurochirurgie
[5]Institut de la Mémoire et de la Maladie d'Alzheimer (IM2A)
[6]Département de Génétique
Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique – Hôpitaux de Paris, 75013 Paris, France
[7]Département de Neuroradiologie, Centre Hospitalier Universitaire de Lille, 59000 Lille, France
*Correspondence: mathias.pessiglione@gmail.com
http://dx.doi.org/10.1016/j.neuron.2012.10.017

SUMMARY

The division of human learning systems into reward and punishment opponent modules is still a debated issue. While the implication of ventral prefrontostriatal circuits in reward-based learning is well established, the neural underpinnings of punishment-based learning remain unclear. To elucidate the causal implication of brain regions that were related to punishment learning in a previous functional neuroimaging study, we tested the effects of brain damage on behavioral performance, using the same task contrasting monetary gains and losses. Cortical and subcortical candidate regions, the anterior insula and dorsal striatum, were assessed in patients presenting brain tumor and Huntington disease, respectively. Both groups exhibited selective impairment of punishment-based learning. Computational modeling suggested complementary roles for these structures: the anterior insula might be involved in learning the negative value of loss-predicting cues, whereas the dorsal striatum might be involved in choosing between those cues so as to avoid the worst.

INTRODUCTION

Learning to avoid potential harms is essential for survival. A substantial part of avoidance learning is based on the experience of punishments following mistakes. Theoretically, punishment-based learning can be modeled with the same computations as reward-based learning. A standard computational solution consists of using prediction errors to update the values on which choices are based (Sutton and Barto, 1998). Biologically, the question of whether reward and punishment learning rely on

a same, common system or on distinct, opponent systems is still debated. A large body of evidence has implicated the ventral prefrontostriatal circuits in encoding reward cues and outcomes (Rutledge et al., 2010; Palminteri et al., 2009a; Hare et al., 2008). Many anatomo-functional models of reward learning share the idea that reward prediction errors (obtained minus expected reward) are encoded in dopamine signals that reinforce corticostriatal synapses (Bar-Gad and Bergman, 2001; Frank et al., 2004; Doya, 2002). The same mechanism could account for punishment learning: dips in dopamine release might weaken approach circuits and/or strengthen avoidance circuits. This is consistent with numerous studies showing that dopamine enhancers improve reward learning, but impair punishment learning in patients with Parkinson's disease (Frank et al., 2004; Bódi et al., 2009; Palminteri et al., 2009b). It has been suggested that another neuromodulator, serotonin, could play an opponent role: it would encode punishment prediction errors (obtained minus expected punishment) so as to reinforce the avoidance pathway (Daw et al., 2002). However, this hypothesis has been challenged by several empirical studies in monkeys and humans (McCabe et al., 2010; Palminteri et al., 2012; Miyazaki et al., 2011).

Beyond neuromodulation, the existence of opponent regions, which would process punishments as the ventral prefrontal cortex and striatum process reward, remains controversial. In humans, fMRI studies of reinforcement learning have yielded inconsistent results. At the cortical level, several candidates for an opponent punishment system have been suggested, among which the anterior insula emerged as particularly prominent. Indeed, the anterior insula was found to represent cues predicting primary punishments, such as electric shocks, fearful pictures, or bad tastes, and these punishments themselves (Büchel et al., 1998; Seymour et al., 2004; Nitschke et al., 2006). These findings have been later extended to more abstract aversive events, such as financial loss or risk (Kuhnen and Knutson, 2005; Samanez-Larkin et al., 2008; Kim et al., 2011, 2006). However, some studies have also found insular activation linked
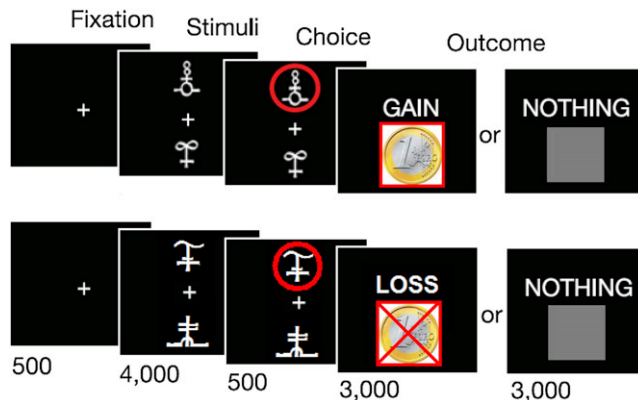
Figure 1. Behavioral Task

Subjects selected either the upper or lower of two abstract visual cues presented on a display screen by pressing ("go") or not ("no go") the space bar of a laptop. They subsequently observed the outcome. In the top examples, the upper cues are chosen ("go" responses), but please note that a given cue appeared in the upper or lower position randomly, so that the motor dimension ("go" versus "no go") was orthogonal to the value dimension ("good" versus "bad"). In the example displayed on the top, the chosen cue is associated with 0.8/0.2 probability of winning 1€/nothing (gain condition, good cue). In the bottom example, the chosen cue is associated with 0.8/0.2 probability of losing 1€/nothing (loss condition, bad cue). Durations of the successive screens are given in milliseconds.

to positive reinforcers and orbitofrontal activation linked to negative reinforcers (O'Doherty et al., 2001; Gottfried and Dolan, 2004; Kirsch et al., 2003). The functional opponency between ventral prefrontal cortex and anterior insula, in learning to predict reward versus punishment, is therefore far from established. At the striatal level, many fMRI studies have reported activations related to primary or secondary reinforcers during instrumental learning (O'Doherty et al., 2003; Galvan et al., 2005; Pessiglione et al., 2008; Daw et al., 2011). Again, some studies supported the idea that the same regions encode both reward and punishments cues or outcomes, whereas other studies argued for a functional dissociation between ventral and dorsal regions (Jensen et al., 2003; Delgado et al., 2000; O'Doherty et al., 2004; Seymour et al., 2007). Thus, while striatum implication in reinforcement learning is indisputable, intrastriatal functional segregation between reward and punishment processing remains to be demonstrated.

Some limitations inherent to fMRI might explain the discrepancies in the literature investigating reward and punishment learning. Because of limited spatial resolution, fMRI activations might confound the activities of neuronal populations encoding distinct, or even opposite, features of the environment. Moreover, the relationship between spiking activity and blood-oxygen-level-dependent signal is not straightforward. In particular, fMRI activation could result from either excitatory or inhibitory signal at the neural level, which may confound punishment and reward encoding. Finally, it remains unclear whether a brain region that activates with reward and deactivates with punishment is involved in reward learning specifically or in both reward and punishment learning. Here we address the existence of an opponent avoidance system by

testing the effect of brain damage on punishment-learning versus reward-learning ability. Showing impaired behavior following brain damage enables conclusions to be made about the causal implication of specific brain regions. This is particularly important for brain areas involved in emotional processing, like the insula, which may represent epiphenomenal reactions that are not causally responsible for producing the behavior.

Another source of confusion comes from the fact that signaling negative values often occur together with implementing inhibition or avoidance behavior. Thus, a brain structure responding to negative cues may not be involved in punishment-based learning, but instead in selecting an action to avoid negative outcome. Here, we use computational modeling to distinguish deficits in reinforcement learning and action selection. Finally, some confusion may have arisen from tasks testing punishment learning in a separate condition and informing subjects that their goal is to avoid punishments. This could shift the frame for outcomes such that not being punished becomes rewarding and hence recruits reward instead of punishment areas. Here we employ a task that mixes reward and punishment learning such that subjects experience both positive and negative outcomes throughout the experiment.

This task (Figure 1) has been previously used for an fMRI study to investigate the effects of dopaminergic medication on instrumental learning (Pessiglione et al., 2006). It involves subjects choosing between two cues to either maximize monetary gains (for reward cues) or minimize monetary losses (for punishment cues). In the previous study, we showed that dopaminergic drugs (levodopa and haloperidol) specifically modulate reward learning, not punishment learning. The aim of the present study is to find brain structures in which lesions would induce the reverse dissociation, impairing punishment learning while leaving reward learning unaffected. Candidates were identified from the previous fMRI results (Figure 2, step 1). Two brain regions were specifically involved in the loss condition: the anterior insula (AI), which was activated in response to both punishment cues and outcomes, and the caudate nucleus (dorsal striatum [DS]), which was only responsive to punishment cues. In contrast, the ventromedial prefrontal cortex (VMPFC) and ventral striatum (VS) were activated in response to reward cues and outcomes. We therefore looked for pathological conditions affecting specifically the AI (not the VMPFC) and the DS (not the VS). For cortical areas, we turned to brain tumors (gliomas) and compared patients with AI damage (INS group) to patients with control lesions elsewhere (LES group). For striatal regions, we turned to Huntington disease and compared presymptomatic patients (PRE group), in whom degeneration is limited to the DS, with symptomatic patients (SYM group), in whom degeneration reaches the VS as well (Douaud et al., 2006; Tabrizi et al., 2009). Two groups of healthy controls (CON) matched to each pathological group of interest (INS and PRE) were also included in the study. All groups performed the exact same instrumental learning task used in the previous fMRI study (Pessiglione et al., 2006) and were tested for an asymmetry between reward- and punishment-based learning.
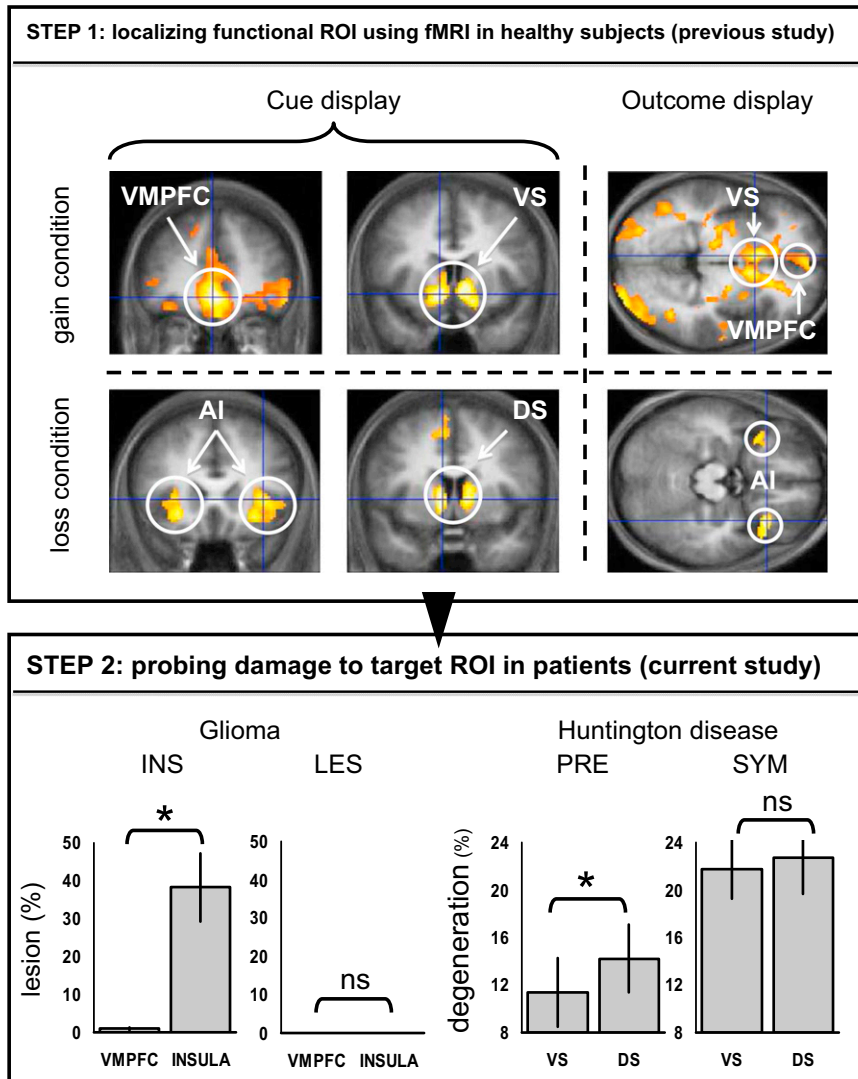
Figure 2. Experimental Design

Functional ROI were identified on the basis of a previous fMRI study of instrumental learning (Pessiglione et al., 2006). Gain-related cues and outcomes activated the VS and VMPFC, whereas loss-related cues and outcomes activated the DS and AI. Patients were selected to assess the effects of specific damage to loss-related regions. For AI, we compared patients with brain lesion over the insula (INS) to patients with lesion elsewhere (LES), whereas for DS, we compared patients at the presymptomatic versus symptomatic stage of Huntington disease (PRE versus SYM HD). Functional activations are shown at a threshold of p < 0.001 uncorrected with a minimum of 60 contiguous voxels and superimposed onto subjects' average anatomical scan. Brain damage is expressed in percentage of volume overlapping with the glioma or in percentage of gray matter loss compared to HD healthy relatives (control group). Error bars represent intersubject SEM. *p < 0.05, paired t test; ns: not significant.

nucleus) were more activated in response to loss versus neutral cue display. At the time of outcome display, losing compared to getting nothing was associated with activations in the bilateral AI and in the anterior cingulate cortex, but not in the DS. These results suggest that, while the AI might be involved in both punishment-based decision and learning phases, the DS might be involved in punishment-based decision only.
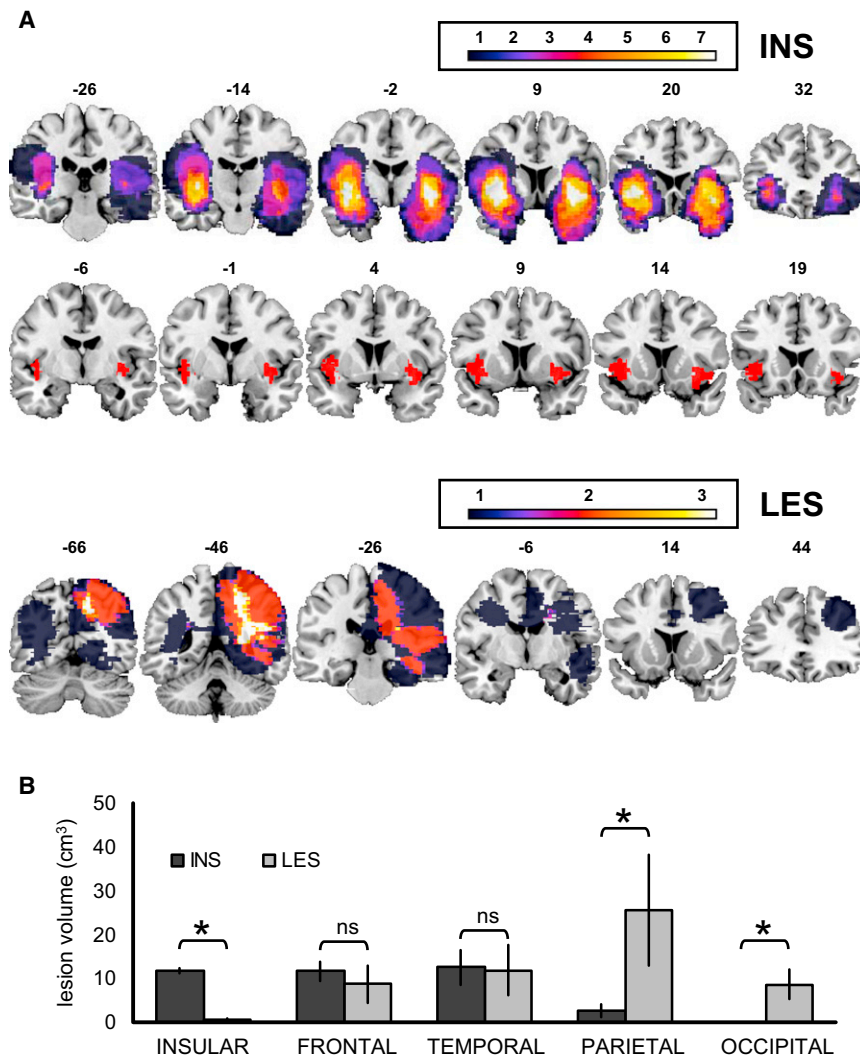
We verified that patient test groups (but not control groups) presented damage to the selected functional ROI (AI and DS; see Figure 2, step 2). Among patients with glioma, the AI, relative to VMPFC, was specifically lesioned in the INS group (38.2 ± 9.1 versus 0.9 ± 0.5 cm^3: t_{13} = 4.1, p < 0.01; paired t test), but remained entirely intact in the LES group. Among patients with Huntington disease (HD), DS gray matter density was preferentially reduced relative to VS in the PRE group (14.2% ± 2.9% versus 11.4% ± 2.8%: t_{13} = 1.9, p < 0.05; paired t test), but not in the SYM group (21.8% ± 2.5% versus 22.7% ± 3.0%; t_{16} = 0.6, p > 0.1; paired t test). These results validate our selection of patient test groups (INS and PRE) as showing preferential damage in punishment-related functional ROI and our selection of patient control groups as presenting intact (LES) or equally atrophic (SYM) reward- and punishment-related areas. We also assessed atrophy in the AI ROI, since insular degeneration has been documented in HD patients (Tabrizi et al., 2009). We found that the AI was unaffected in PRE patients (−0.2% ± 3.8%; t_{13} = 0.5, p > 0.1, paired t test), but significantly atrophic in SYM patients (8.2% ± 3.3%; t_{16} = 2.5, p < 0.05, paired t test). We hereafter provide more details about the anatomical localization of brain damage

RESULTS

Brain Damage Delineation

Cortical and striatal regions of interest (ROI) were based on a re-analysis of previous fMRI data (Pessiglione et al., 2006), focusing on the placebo group (n = 13) to avoid biases due to pharmacological manipulation. The different cues and outcomes (gain, neutral, and loss) were modeled with separate regressors in a general linear model (GLM). Regression coefficients (betas) were then contrasted and tested for significance at the group level (with a voxel threshold of p < 0.001 uncorrected and a cluster threshold of p < 0.05 after family-wise error (FWE) correction for multiple comparisons). Gain-predicting cues, compared to neutral or loss-predicting cues, elicited activity in the VMPFC, VS, and posterior cingulate cortex. The same regions were also activated at the outcome onset when winning compared to getting nothing. These results support the implication of ventral prefrontostriatal circuitry in reward-based decision and learning. The bilateral AI and bilateral DS (head of caudate

A

B

threshold of 60 contiguous voxels) as for the functional activation analysis described above. Carriers of the HD mutation (>36 CAG repeats in the HTT gene) were split into PRE (n = 14) and SYM (n = 17) groups, depending on whether their motor symptoms, evaluated by the Unified Huntington's Disease Rating Scale (UHDRS) scores, were smaller or bigger than 5/124. A group of healthy relatives (CON, n = 14) was also included in the VBM analysis. An ANOVA was performed on individual gray matter density maps with group (CON, PRE, and SYM) as the main factor of interest. The main effect of group was a significant atrophy in the basal ganglia (Figure 4A), more precisely in the bilateral anterior parts of the putamen, caudate, and pallidum, as well as in the amygdala and thalamus. Within these regions, we performed post hoc t tests to compare groups two by two. Compared to the CON group, the PRE group showed a significant atrophy specific to the bilateral caudate (Figure 4A). Compared to the PRE group, the SYM group showed a significant atrophy in the ventral parts of the anterior putamen and pallidum, as well as in the amygdala and thalamus (Figure 4A). The observed pattern of neurodegeneration is therefore consistent with previous studies reporting a dorsoventral gradient of striatal gray matter loss in HD (Douaud et al., 2006; Tabrizi et al., 2009). Thus our whole brain analysis confirmed that the dorsoventral gradient is pronounced in presymptomatic, but attenuated in more advanced stages of the disease. This observation was further supported by direct comparison between anatomically defined ROI (Figure 4B): the caudate nucleus was more atrophic than the ventral striatum in PRE patients (15.2% ± 2.9% versus 11.0% ± 2.8%; $t_{13} = 2.5$, p < 0.05, paired t test), but not in SYM patients (23.7% ± 3.2% versus 21.8% ± 2.7%; $t_{16} = 1.0$, p > 0. 1, paired t test).

in the different patient groups, independently of functional activations.

Regarding patients with brain tumors (gliomas), we computed an overlap map of individual lesions normalized onto an anatomical template (Figure 3A). Patients were split into the INS (n = 14) and LES (n = 9) groups, depending on whether their lesions affected the insula or not. In the INS group, the maximum of overlap (n = 7 for each hemisphere) specifically covered the insular lobe. Note that, because lesions were unilateral, the greatest possible overlap with the bilateral functional AI ROI is 50%. Other areas were also damaged in the frontal (11.7 ± 2.2 cm³), temporal (12.5 ± 4.0 cm³), and parietal (2.7 ± 1.5 cm³) lobe. However, for each lobe, the volume of these extrainsular lesions in the INS group was similar or lesser than in the LES group (Figure 3B). Thus the only brain area that was more damaged in the INS compared to the LES group was the insula (11.9 ± 0.6 versus 0.6 ± 0.4 cm³, $t_{20} = 12.9$, p < 0.001, two-sample t test).
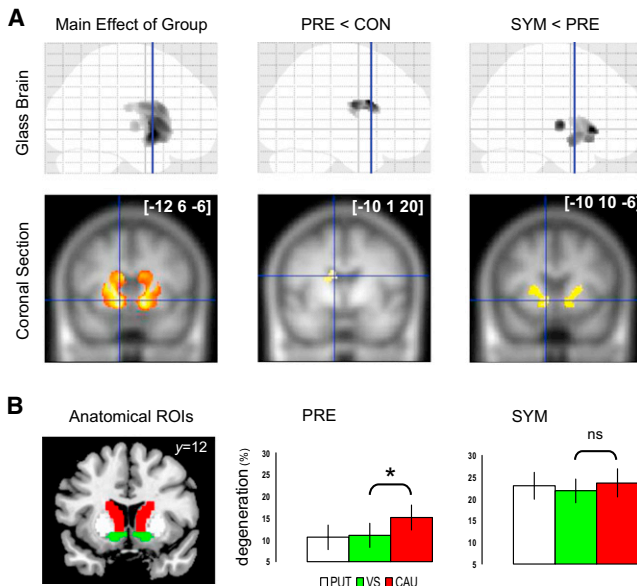
Regarding patients with HD, we used voxel-based morphometry (VBM) analysis to quantify cerebral atrophy, using the same statistical threshold (p < 0.001 uncorrected with an extent

Behavioral Analysis
All subjects were able to learn over the 30 trials of a learning session the correct response, which was choosing the most

Figure 4. Atrophy Delineation in Patients with Huntington Disease
(A) Statistical parametric maps represent the main effect of group (left column), the comparison of PRE to CON patients (middle column), and the comparison of SYM to PRE patients. Areas colored in a gray-to-black gradient on sagittal glass brains and in a red-to-white gradient on coronal slices showed significant gray matter reduction (p < 0.001 uncorrected with a minimum of 60 contiguous voxels). [x y z] coordinates of maxima refer to the MNI space.
(B) ROI analysis of HD patients atrophy. Left: ROI are illustrated on a coronal slice. Right: percentage of gray matter loss compared to HD relatives (control group) for each group and ROI. Error bars represent intersubject SEM. *p < 0.05, paired t test; ns: not significant.

rewarding cue in the gain condition and avoiding the most punishing cue in the loss condition (Figure 5). The difference between average percentage of correct choices in the gain and loss conditions, which we termed the reward bias, was compared between groups using ANOVA (Figure 6).

Testing the impact of glioma, we found a significant group effect on the reward bias ($F_{2,40} = 4.7$, p < 0.05). Post hoc comparisons using two-sample t tests showed that the reward bias was higher in the INS group compared to both CON ($t_{32} = 3.0$, p < 0.01) and LES ($t_{21} = 2.0$, p < 0.05) groups. In fact, paired t tests demonstrated a significant reward bias in the INS group ($t_{13} = 4.5$, p < 0.001), but not in the CON and LES groups ($t_{19} = 1.2$, p > 0.1 and $t_8 = 1.0$, p > 0.1). The group effect on the reward bias was driven by a significant group effect on punishment learning ($F_{2,40} = 3.2$; p < 0.05), contrasting with an absence of group effect on reward learning ($F_{2,40} = 0.4$; p > 0.5). Post hoc comparisons showed that punishment learning was significantly impaired in INS patients relative to both CON ($t_{32} = 2.1$, p < 0.05) and LES patients ($t_{21} = 1.9$, p < 0.05), with no difference between CON and LES groups ($t_{27} = 0.1$, p > 0.5). To control for lateralization of brain damage, we compared punishment-learning performance between right- and left-lesioned patients: there was no significant difference ($t_{12} = 0.1$, p > 0.5). To control for size, we regressed punishment-learning performance against lesion volume: there was no significant correlation ($R^2 = 0.03$, p > 0.5).
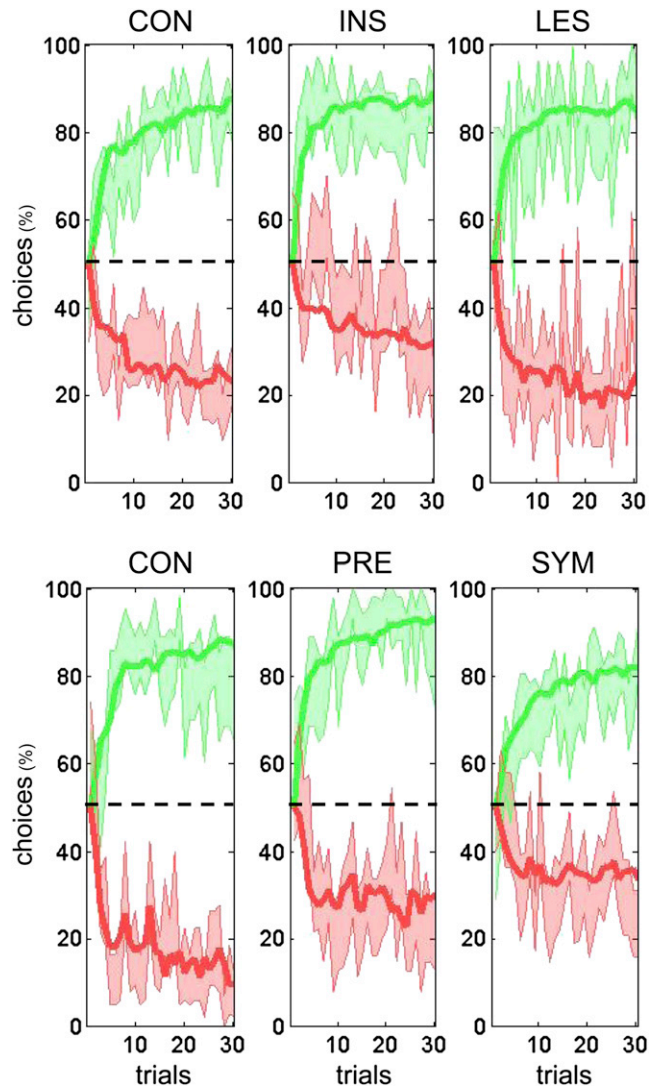
Figure 5. Learning Curves
Trial-by-trial average choices are illustrated for the two conditions (green: gain; red: loss) in the different groups (healthy controls [CON], INS, LES, PRE, and SYM). Colored areas represent observed choices (mean ± SEM); bold lines represent modeled choices.

Testing the impact of HD, the ANOVA performed on the reward bias showed a significant group effect ($F_{2,42} = 4.6$; p < 0.05). Post hoc two-sample t tests showed that the reward bias was higher in the PRE group compared to both CON ($t_{26} = 3.4$, p < 0.01) and SYM ($t_{26} = 1.7$, p < 0.05) groups. Paired t tests demonstrated a significant reward bias in the PRE group ($t_{13} = 4.8$, p < 0.001), but not in the CON and SYM groups ($t_{13} = 0.6$, p > 0.1 and $t_{16} = 1.3$, p > 0.1). Again, the reward bias effect was driven by a significant group effect on punishment learning ($F_{2,42} = 3.8$; p < 0.05), contrasting with an absence of significant group effect on reward learning ($F_{2,42} = 2.1$; p > 0.1). Compared to CON patients, post hoc t tests showed a significant reduction of punishment learning in both PRE ($t_{26} = 1.8$, p < 0.05) and SYM ($t_{29} = 2.7$, p < 0.01) patients, but no significant difference
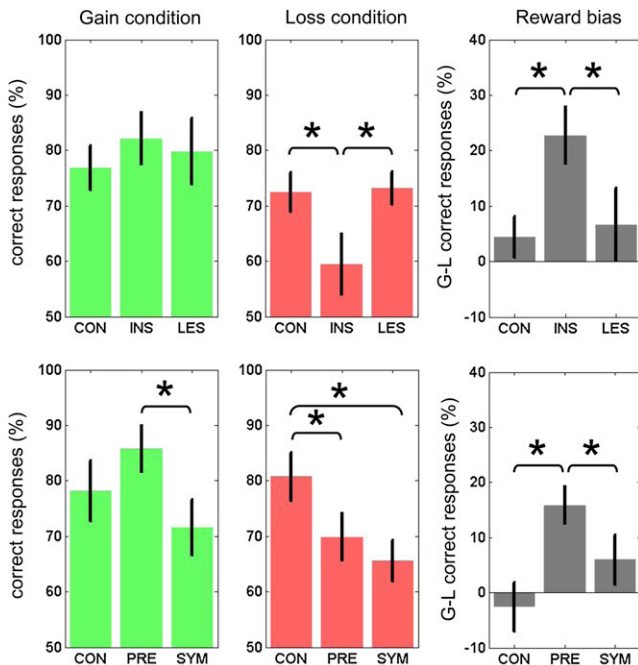
Figure 6. Behavioral Analysis
Behavioral performance (% of correct responses) is illustrated for the two conditions (green: gain; red: loss) in the different groups (CON, INS, LES, PRE, and SYM). The reward bias is the difference between gain and loss conditions. Error bars represent intersubject SEM. *p < 0.05, two-sample t test.

between PRE and SYM groups ($t_{29} = 0.8$, p > 0.1). However, SYM patients showed a significant reduction in reward learning compared to PRE patients ($t_{29} = 1.8$, p < 0.05) or compared to PRE and CON groups pooled together ($t_{29} = 1.8$, p < 0.05). This difference was still significant when including treatment as a covariate and therefore was not due to neuroleptics impeding reward learning. There was a trend toward reward learning impairment with neuroleptics, but this was not significant (medicated: 69.7% ± 6.3%, unmedicated: 75.0% ± 9.1% of correct responses; $t_{13} = 0.5$, p > 0.1, two-sample t test). We also tested direct Pearson's correlation of learning performance with gray matter density extracted for each patient from group-level caudate ROI (i.e., from the significant cluster obtained in PRE < CON contrast; see Figure 4A). The correlation was marginally significant for the punishment condition ($R^2 = 0.41$; p < 0.07), but not for the reward condition ($R^2 = 0.15$; p > 0.2).

In summary, we found an asymmetry in favor of reward-based relative to punishment-based learning specifically in patients with anterior insula lesion (INS group) and in patients with dorsal striatum atrophy (PRE group).

Computational Analysis

The observed deficits in punishment learning needed further characterization, as obviously the average percentage of correct responses does not assess learning dynamics. We analyzed learning dynamics in more details by fitting a standard Q-learning model (Sutton and Barto, 1998) to the observed choices (Figure 5). The model combines the Rescorla–Wagner learning

rule, which updates chosen option values in proportion to reward prediction errors, and a softmax decision rule, which estimates choice probability as a sigmoid function of the difference between the two option values. Fitting the model to learning curves means adjusting the free parameters to maximize the likelihood of the observed choices. This was done separately for the gain and loss conditions in each subject. Then the adjusted free parameters, namely the learning rate (α), choice randomness (β), and reinforcement magnitude (R), were systematically tested for group effect using ANOVA (Figure 7).

Regarding the glioma groups, there was no significant effect in the gain condition (α_G: $F_{2,40} = 0.3$, p > 0.5; β_G: $F_{2,40} = 0.2$, p > 0.5; R_G: $F_{2,40} = 0.4$, p > 0.5). In the loss condition, we found a significant group effect for the reinforcement magnitude (R_L: $F_{2,40} = 3.2$; p < 0.05), but not for the learning rate (α_L: $F_{2,40} = 0.0$, p > 0.5) or choice randomness (β_L: $F_{2,40} = 0.6$, p > 0.5). Post hoc comparisons using two-sample t tests found that, in the INS group, the R_L was significantly reduced compared to CON ($t_{32} = 2.3$, p < 0.05) and LES ($t_{21} = 2.3$, p < 0.05) groups.

Regarding HD patients, the same ANOVA revealed no significant group effect for any parameter estimate in the gain condition (α_G: $F_{2,42} = 0.1$, p > 0.5; β_G: $F_{2,42} = 0.1$, p > 0.5; R_G: $F_{2,42} = 1.8$, p > 0.1). In the loss condition, the only significant effect was found for choice randomness (β_L: $F_{2,42} = 4.2$; p < 0.05), not for learning rate (α_L: $F_{2,42} = 0.6$; p > 0.5) or reinforcement magnitude (R_L: $F_{2,42} = 1.4$; p > 0.1). Post hoc t tests showed that, relative to the CON group, β_L was significantly higher in both PRE and SYM groups, ($t_{26} = 1.8$, p < 0.05 and $t_{26} = 2.7$, p < 0.01). In the gain condition, the only significant difference was a higher R_G in the PRE compared to the SYM group ($t_{29} = 1.7$, p < 0.05).

In summary, the computational analysis indicated that the observed punishment-based learning deficit was specifically captured by a lower reinforcement magnitude (R_L) parameter in the INS group and by a higher choice randomness (β_L) parameter in the PRE group. In order to statistically assess that the affected parameter depended on the site of brain damage, we ran an ANOVA with group (INS and PRE) as a between-subject factor and effect (reduction in R_L and $1/\beta_L$ relative to controls) as a within-subject factor. Crucially, we found a significant group by effect interaction ($F_{1,26} = 4.4$, p < 0.05), supporting the idea that different computational parameters were affected in the INS and PRE groups.

DISCUSSION

Here we tested the performance of brain-damaged patients with an instrumental learning task that involves both learning option values and choosing the best option. Behavioral results indicate that both damage to the AI and degeneration of the DS specifically impair punishment avoidance, leaving reward obtainment unaffected. Computational analyses further suggest that AI damage affects the learning process (updating punishment values), whereas DS damage affects the choice process (avoiding the worst option).

The instrumental learning task used to demonstrate this dissociation has several advantages. A first advantage is that money offers comparable counterparts for reward and punishments,
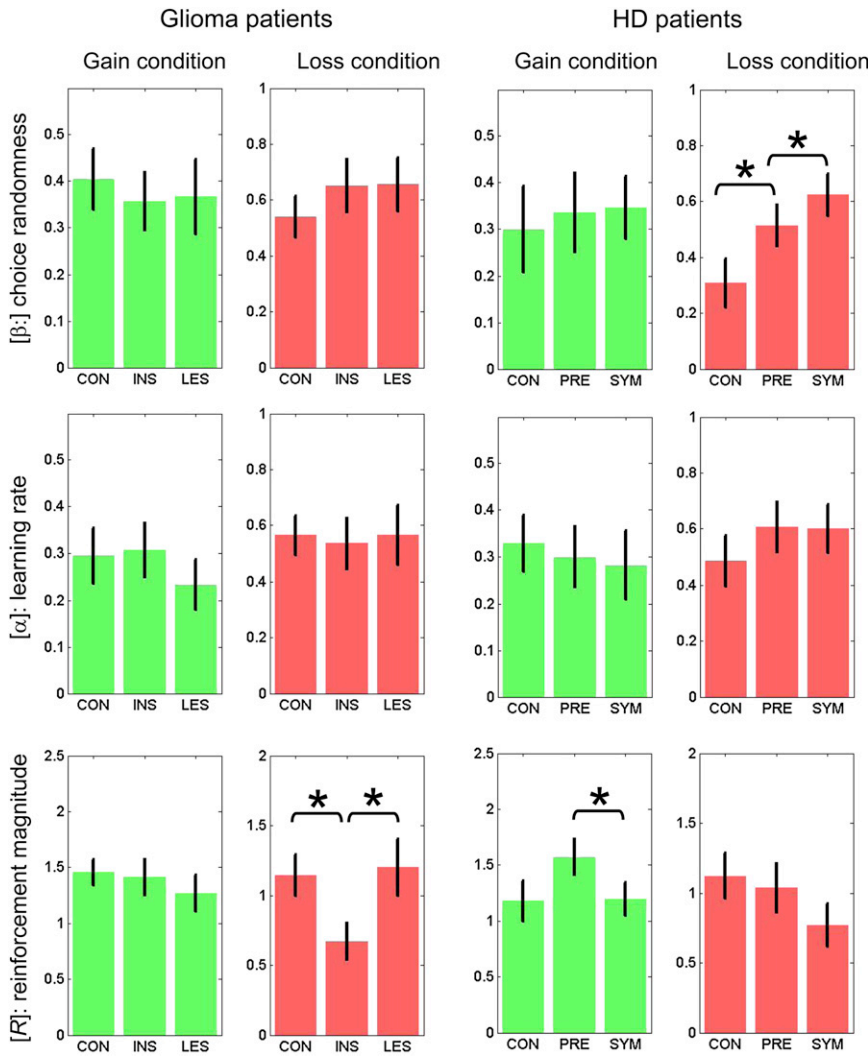
other option in addition to avoid choosing the worst one. Also, in reward learning, subjects get more reinforcement as soon as they select the correct response, whereas in punishment learning they get less reinforcement. This would support the idea that punishment avoidance is more difficult and hence more sensitive to brain damage. However, we found the reverse dissociation, meaning a selective effect on reward learning, in the exact same task with dopaminergic drugs (Pessiglione et al., 2006). Thus a difference in sensitivity is unlikely to explain the selective effects of AI and DS damage on punishment learning. It remains nonetheless possible that, once subjects have learned the valence of symbols, they reframe their expectations such that neutral outcomes become punishing in the gain condition and rewarding in the loss condition. However, this should have blurred the difference between reward and punishment conditions and therefore contributed to diminish, not induce, the asymmetry that we observed in our data.

The same instrumental learning task was used in a previous fMRI study that we reanalyzed to identify candidate regions (AI and DS) for underpinning punishment-based learning and avoidance. We benefited from the rare opportunity to test damage to these ROI in hospitalized patients. Indeed, the Pitié-Salpêtrière hospital contains a neurosurgery ward capable of removing glioma located around the anterior insula, which presents difficulties due to the proximity of Broca's area (Jones et al., 2010). Also, our hospital is a national reference center for Huntington disease that participates in the international multicentric longitudinal study Track-HD (Tabrizi et al., 2009). To our knowledge, avoidance learning ability had never been investigated in patients with insular lesion (INS) nor HD. We checked that tumoral masses overlapped with functional AI in INS patients and that neural atrophy overlapped with functional DS in presymptomatic HD patients. Naturally this overlap was only partial, meaning that functional ROI were not entirely destroyed and that other regions were also attained, which may complicate the attribution of punishment avoidance deficits to selective anatomical structures. Insular damage (INS patients) was unilateral, covering almost 40% of the functional AI, which was bilateral. Striatal degeneration (PRE patients) was bilateral, but

contrary to the reinforcements used in animal conditioning, such as fruit juice and air puff (Ravel et al., 2003; Joshua et al., 2008; Morrison and Salzman, 2009). However, the well-known phenomenon of loss aversion (Tversky and Kahneman, 1992; Tom et al., 2007) suggests that financial punishment may have more impact than financial reward of the same amount. This effect would go against our finding that punishment learning was deficient in patients. Reciprocally, it could be argued that losses had less impact because patients were not playing with their own real money. It is important to note here that double dissociations between outcome valence and dopaminergic medication have been obtained with virtual money or even with points (Frank et al., 2004; Bódi et al., 2009; Palminteri et al., 2009b). This suggests that instrumental learning performance is sensitive enough to virtual gains and losses, even if real money might elicit stronger responses in some subjects. Another advantage of the task is that reward and punishment conditions are matched in difficulty, as the same probabilistic contingencies were to be learned. One may nonetheless argue that punishment avoidance involves an extra step, since subjects must select the

limited to 15% of the functional DS. In both cases, it remains difficult to state what proportion of the ROI remained truly functional. The fact that we did observe the expected deficits suggests the ROI were significantly impaired, even if not entirely. The presence of brain damage outside the functional ROI raises the question of specificity, which we addressed by including control pathological conditions. We verified that, in both groups of interest (INS and PRE), functional punishment-related regions were more affected than functional reward-related regions, namely the ventromedial prefrontal cortex and ventral striatum. These groups therefore allowed testing the existence of opponent structures at both cortical level with glioma and subcortical level with HD. Each group of interest was compared both to healthy controls and to patients who presented similar lesions, except that punishment-related ROI were not preferentially affected. Thus, the observed deficits can be attributed to punishment-related ROI, perhaps not specifically to the clusters activated in fMRI analyses, but at least to anatomically defined AI and DS.

Computational analyses revealed distinct deficits in INS and PRE patients. The flattened punishment-learning curve following AI damage in INS patients was specifically captured by a reduced reinforcement magnitude in the loss condition. This means that not only learning was slowed down, but also the asymptotic plateau was lower. It can be distinguished from a change in learning rate, which would only affect how fast the plateau is reached. This computational result suggests that AI damage attenuated not only signaling of aversive outcomes, but also signaling of aversive cues. This is consistent with our previous neuroimaging findings that the AI was responsive to both aversive outcome display (during learning period) and aversive cue display (during choice period). It is also more generally consistent with a number of neuroimaging studies that implicated the anterior insula in signaling aversive events (Büchel et al., 1998; Seymour et al., 2004; Nitschke et al., 2006; Samanez-Larkin et al., 2008; Kim et al., 2006, 2011). In our original fMRI analysis, as in some other fMRI studies (Kim et al., 2006; Seymour et al., 2004; Pessiglione et al., 2006), the AI was found to encode more precisely punishment prediction errors (received minus expected punishments) at the time of outcome. This is also consistent with the present computational result, as reducing reinforcement magnitude in the loss condition is one way to diminish punishment prediction error signal. Thus signaling punishment prediction error following outcome might be the computational mechanism by which the AI causally impacts negative value learning. At the time of choice, the AI might signal cue negative value (i.e., punishment prediction), which could drive avoidance behavior. This is in line with theories proposing that brain areas involved in somatic affective representations are causally responsible for making a choice (Jones et al., 2010; Naqvi and Bechara, 2009; Craig, 2003).

The flattened punishment-learning curves following DS preferential atrophy in presymptomatic HD patients was specifically captured by a higher choice randomness. Contrary to reinforcement magnitude and learning rate, this parameter impacts the choice, not the learning process. This is consistent with our fMRI finding that the DS was active at punishment cue display (during choice period), but not at outcome display (during

learning period). It accords well with the idea that the DS is the "actor" part of the striatum, the "critic" part being more ventral (O'Doherty et al., 2004; Atallah et al., 2007). Indeed, the transition from presymptomatic to symptomatic HD, which was characterized by degeneration extending to the VS, was captured by a lower reinforcement magnitude in the gain condition. Thus the VS, which is closely linked to the VMPFC, would play a role similar to that of the insula, but for learning positive instead of negative values. This is in line with studies implicating the VS and VMPFC in encoding both reward predictions at cue display and reward prediction errors at outcome display (Rutledge et al., 2010; Palminteri et al., 2009a; Hare et al., 2008). However, interpreting the specific role of the DS in choosing between aversive cues remains speculative. The link with choice randomness might suggest that the DS is involved in comparing negative value estimates or in integrating the precision of these estimates, or in adjusting the balance between exploration and exploitation. Another possibility is that the DS is specifically involved in avoidance behavior, i.e., in inhibiting the selection of the worst option and facilitating the selection of alternatives. This interpretation is endorsed by the observation that input connections to the caudate head come from dorsal prefrontal structures, which have been implicated in inhibitory and executive processes (Draganski et al., 2008; Haber, 2003; Postuma and Dagher, 2006).

In conclusion, we found evidence that the AI and DS are causally implicated in punishment-based avoidance learning, but for different reasons. The AI might participate by signaling punishment magnitude, in accordance with its involvement in negative affective reactions, whereas the DS might participate by implementing avoidance choices, in accordance with its involvement in executive processes. These findings suggest the existence of a distinct punishment system underpinning avoidance learning, just as the reward system underpins approach learning. However, we only tested two candidate regions here; further research is needed to circumscribe this putative punishment system. Two promising candidates are the amygdala, which has been involved in encoding negative emotions, including when losing money (Yacubian et al., 2006; De Martino et al., 2010; Schlund and Cataldo, 2010), and the anterior cingulate cortex, which is frequently coactivated with the AI when subjects experience negative affective states, also including monetary losses (Blair et al., 2006; Petrovic et al., 2008; Kahnt et al., 2009).

EXPERIMENTAL PROCEDURES

Behavioral Task

We employed here the same task as that previously used for an fMRI study (Pessiglione et al., 2006), except that pounds were changed into euros, that English was translated into French, and that, in order to shorten the experiment, subjects performed only two test sessions after one full training session. Subjects were provided with written instructions, which were reformulated orally when necessary, asking them to try and maximize their financial payoff (see Figure 1; Supplemental Information available online). Each session was an independent task containing three new pairs of cues to be learned. Cues were abstract visual stimuli taken from the Agathodaimon alphabet. Each pair of cues was presented 30 times for a total of 90 trials. The three cue pairs corresponded to the three conditions (gain, neutral, and loss), which were respectively associated with different pairs of outcomes (winning 1€ versus nothing, looking at 1€ versus nothing, and losing 1€ versus nothing). Within each pair, the two cues were associated to the two possible outcomes with reciprocal

probabilities (0.8/0.2 and 0.2/0.8). Note that there was no financial outcome associated with the neutral cue, for which the euro coin could be looked at, but not won or lost. On each trial, one pair was randomly presented and the two cues were displayed on a computer screen above and below a central fixation cross, their relative position being counterbalanced across trials. The subject was required to choose the upper stimulus by pressing the space bar ("go" response) or the lower stimulus by retaining from pressing any button ("no go" response) within a 4 s delay. Please note that, since the position on screen was counterbalanced, response (go versus no go) and value (good versus bad cue) were orthogonal. After the 4 s delay, the chosen cue was circled in red and then the outcome (either "nothing," "gain," "look," or "loss") was displayed on the screen. In order to win money, subjects had to learn by trial and error the cue–outcome associations, so as to choose the most rewarding cue in the gain condition and the less punishing cue in the loss condition.

Regions of Interest

To identify ROI, we reanalyzed, using SPM8 software, the fMRI data acquired for a previous published study (Pessiglione et al., 2006) that investigated the effects of dopaminergic drugs on instrumental learning. Please refer to this publication for details about image acquisition and preprocessing. To avoid drug-induced biases, we only analyzed here the data from the 13 subjects who were assigned to the placebo group in the original study. Individual time-series were analyzed using a general linear model that included separate regressors for the different events and conditions. Two events were considered in every trial, cue, and outcome onset, which were modeled with a delta function. There were three different conditions for cues: gain, neutral, and loss. There were six different types of outcomes: winning £1 or getting nothing in the gain condition, looking at £1 or getting nothing in the neutral condition, and losing £1 or getting nothing in the loss condition. We then computed between-cues and between-outcomes linear contrasts to identify brain regions specifically implicated in gain and loss processing. Individual contrasts were taken to a group-level random-effect analysis using one-sample t tests. Activations reported here survived a cluster-forming threshold of p < 0.001 (uncorrected), with an extent threshold of 60 contiguous voxels to ensure significance of p < 0.05 after family-wise error correction for multiple comparisons over the whole brain.

To verify that our patients constituted valid models of lesions in the targeted ROI, we built mask images by taking the intersection between functional clusters significantly activated in the relevant contrasts and anatomical areas delineated with MARINA software. The VMPFC mask was defined as the intersection between the contrast of gain-predicting versus loss-predicting cues and an anatomical template composed of the orbital parts of the superior, middle, and inferior frontal gyri, as well as the gyrus rectus, olfactory cortex, and anterior cingulate cortex (all bilateral). The VS mask was defined as the intersection between the contrast of gain-predicting versus neutral cues, the contrast of £1 versus £0 outcomes, and an anatomical template, including the bilateral putamen and caudate nuclei. The DS mask was defined as the intersection between the contrast of loss-predicting versus neutral cues and the same anatomical template for the bilateral putamen and caudate nuclei. The AI mask was defined as the intersection between the contrast of loss-predicting versus neutral cues, the contrast of the −£1 versus £0 outcomes, and an anatomical template, including the bilateral insula.

Subjects

The study was approved by the ethical committee of the Pitié-Salpêtrière Hospital, where the study took place. In total we tested 88 subjects: 34 controls and 54 patients (23 with brain tumors and 31 with Huntington disease). Most healthy controls were relatives accompanying patients to the hospital. All subjects gave written informed consent prior to inclusion in the study. They were not paid for their voluntary participation and were told that the money won in the task was purely virtual. Previous studies have shown that using real money is not mandatory to obtain robust motivational or conditioning effects (Frank et al., 2004; Palminteri et al., 2009b). In our case, using real money would be unethical, since it would mean penalizing patients for their handicap. See Table S1 for detailed clinical and demographical data.

Brain Tumors

Patients were hospitalized either for a biopsy, in order to determine the nature of the tumor (n = 5), or for the surgical ablation of the tumor (n = 18). Tumors were gliomas in all patients but one, in whom the tumor was metastatic. The precise glioma types were (grades are given following the World Health Organization classification): 12 oligoastrocytomas (grade 2: n = 7; grade 3: n = 5), three oligodendrogliomas (all grade 2), two astrocytomas (grade 2 for one and grade 3 for the other), one pilocytic astrocytoma (grade 1), and two glioblastomas (grade 4). For two gliomas, the precise type could not be determined (one was grade 2 and the other grade 3). A majority of patients (15/23) were under preventive antiepileptic medication because of a history of tumor-related seizure. No patient was taking any medication interfering with the dopaminergic system, such as neuroleptics. Patients were tested 29 ± 13 (mean ± SEM) months after the onset of clinical symptoms and 24 ± 12 months after the MRI or computed tomography scan that had confirmed the diagnosis of tumoral mass present in the brain. Patients were split according to whether the lesion overlapped with the insula (INS group: n = 14) or not (LES group: n = 9). Tumor etiology was globally matched between the two groups, with similar grades (INS: 2.4 ± 0.2; LES: 2.1 ± 0.2; p > 0.3, t test) and a similar proportion of oligoastrocytomas (INS: 8/14; LES: 4/9; p > 0.4, chi2-test). We also checked that lesion sizes were comparable between the two groups (INS: 76.6 ± 10.8; LES: 92.0 ± 22.0; p > 0.5, t test). A cohort of healthy subjects was also included (CON group; n = 20). These subjects were matched to INS patients in age (CON: 43.6 ± 2.8; INS: 46.7 ± 3.9; p > 0.5, t test), gender (CON: 12/8; INS: 9/5; p > 0.7, chi2-test), and handedness (CON: 16/4; INS: 11/3; p > 0.9, chi2-test). There was no cognitive impairment in the INS group, as indicated by the normal Mini-Mental State (MMS) score (29.3 ± 0.6). INS patients were not depressed (Hospital Anxiety and Depression [HAD] depression score: 4.9 ± 0.7), but moderately anxious (HAD anxiety score: 8.2 ± 1.2). Unfortunately, the MMS and HAD scores were only collected for a minority of LES patients (4/9), in whom they were similar to those obtained in the INS group (MMS: 30.0 ± 0.0; HAD depression: 4.5 ± 2.4; HAD anxiety: 10.3 ± 2.9).

Delineation of Brain Lesion

All lesioned patients but one had a high-definition three-dimensional anatomical T1 MRI scan and a fluid attenuated inversion recovery T2 MRI scan. The scans were acquired on average 39.6 ± 23.6 days before the experiment. Based on both T1 and T2 scans, the tumoral masses were manually segmented on the native anatomical space using MRIcro (http://www.cabiatl.com). The T1 scans were normalized to an anatomical template with the Statistical Parametric Mapping software (SPM8: http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) running on Matlab (http://www.mathworks.com). The resulting images were carefully checked one by one to ensure that the lesion did not perturb the normalization process. The same transformations computed to normalize T1 scans were then applied to the corresponding lesion images. Overlap maps were built by summing the lesion images separately for the two patient groups (INS and LES). To analyze the spatial distribution of lesions, we built anatomical masks of the insular, frontal, parietal, temporal, and occipital lobes based on the automatic anatomic labeling atlas (AAL), as implemented by the MARINA software (http://www.bion.de). We quantified the volume of the intersections between individual lesion images and every anatomical mask. We then compared these volumes between INS and LES patients using two-sample t tests. To verify that the glioma selectively impacted our functional ROI, we calculated the percentage of voxels within the AI and VMPFC masks overlapping with the lesion images and compared this overlap between ROI in each group (INS and LES) using paired t tests.

Huntington Disease

We included 45 subjects participating in the Paris site of the Track-HD study, a multicentric research protocol that has been designed to study the early stages of HD (Tabrizi et al., 2009). Among these subjects, 31 were carriers of the mutation leading to HD (abnormal CAG expansion in the HTT gene). These patients were split into presymptomatic (PRE, n = 14) and symptomatic (SYM, n = 17) groups, depending on their scores in the UHDRS, with a cut-off at 5/124, as previously reported (Tabrizi et al., 2009). The mean estimated duration to onset in the PRE group was 9.4 years, and the mean duration from onset in the SYM group was 5.2 years. Note that the SYM group was still in an early stage of HD. The other 14 subjects were not carriers of the HD mutation and

therefore considered as healthy controls (CON, n = 14). They were either the partners or the siblings of other (nonincluded) HD patients. Control subjects were matched to presymptomatic patients for demographic variables, such as age (CON: PRE: 46.4 ± 3.1; p > 0.3, t test), gender (CON: 8/6; PRE: 7/7, p > 0.5, chi2-test), and handedness (CON: 13/1; PRE: 13/1), as well as for clinical variables, such as the UHDRS score (CON: 1.4 ± 0.3; PRE: 2.1 ± 0.4; p > 0.1, t test) and the MMS score (CON: 29.6 ± 0.2; PRE: 29.7 ± 0.2; p > 0.7, t test). Symptomatic patients differed from presymptomatic patients on UHDRS scores (PRE: 2.1 ± 0.4; SYM: 16.9 ± 2.1; p < 0.001, t test). Among presymptomatic patients, one was taking anxiolytic treatment at the moment of the test and one was under neuroprotecting preventive therapy. No presymptomatic patient was taking any medication interfering with dopaminergic functions. Among symptomatic patients, 11/17 were taking neuroleptics and 9/17 anxiolytics.

Delineation of Brain Atrophy

All subjects (both HD patients and their relatives) included in the Track-HD protocol had a three-dimensional anatomical T1 MRI scan. For every step of the VBM analysis, we used the diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL) toolbox implemented in SPM8 software. Following the standard procedure outlined in the VBM tutorial (http://www.fil.ion.ucl.ac.uk/~john/misc/VBMclass10.pdf), the images were first segmented in the native space into six classes of tissues: gray matter (GM), white matter (WM), cerebral spinal fluid (CSF), skull, soft tissue outside the brain, and a last class accounting for air and remaining signal outside the head. Importantly, this first step generated a roughly (via a rigid-body transformation) aligned GM and WM image for every subject. Both GM and WM images were then warped to an iteratively improved template using nonlinear registration in DARTEL. This step produced the final DARTEL template and the corresponding deformation fields used to match each gray matter image to this template. Finally, the DARTEL template was registered to the Montreal Neurological Institute (MNI) space using affine transformation. This transformation and the DARTEL flow-fields were used to warp the GM images in a way that preserved their local tissue volumes. A Gaussian kernel of 8 mm full-width at half-maximum was then applied for spatial smoothing.

The individual GM images were entered in a full factorial design analysis with group as the main factor. The total intracranial volume was also entered in the statistical model as a covariate to control for confounding effects of brain size. Since our groups were matched regarding demographic variables, these were not included in the model. We first analyzed the main effect of group using F-contrast. Significance threshold was set at p < 0.001 (uncorrected) with an extent threshold of 60 contiguous voxels. Significant clusters in this main group effect were pooled to build a mask for subsequent group comparisons (CON versus PRE and PRE versus SYM) using two-sample t tests. Anatomical labeling of significant clusters was obtained by superimposing the statistical parametric maps to the AAL atlas implemented in MRIcro software. To examine how atrophy impacted our striatal ROI (VS and DS), we extracted the percentage of gray matter in each group and compared the loss of gray matter (relative to HD controls) between the two regions (VS and DS) in each patient group (PRE and SYM) using paired t test. We also defined three anatomical a priori ROI to examine the degeneration pattern over the VS, caudate, and putamen nuclei. These ROIs were manually segmented using MRIcro software on the single subject T1 template of SPM8 software.

Data Analysis

Performance in the first training session was significantly poorer than in the two test sessions, whatever the group. This first session was therefore considered as a practice and not analyzed further. However, the main results (significant group by condition interactions) were also observed when including this first session in the analysis. The neutral pair was introduced for fMRI contrasts, but was discarded for behavioral data analysis. Results therefore come from the two other sessions, each containing 30 trials for each condition (gain or loss). Percentage of go responses was not further considered, as it was similar in all groups and not different from 50%, which comes from the correct cue being on top of the screen in half the trials. We extracted three dependent variables, which we termed gain learning, loss learning, and reward bias. Gain and loss learning were the average percentage of correct choices in the gain and loss condition. Reward bias was the difference between gain and loss learning.

To test the effects on brain damage, we compared the reward bias between groups with an ANOVA. Note that testing group effect on the reward bias is formally equivalent to testing a group by condition interaction. Significant effects were further analyzed with post hoc between-group comparisons separately performed on the dependent variables using two-sample t tests.

To further investigate which process was affected by brain damage, we fitted the learning curves with a computational model. We used the same standard Q-learning algorithm that was employed to capture the effects of dopaminergic drugs in a previous fMRI study (Pessiglione et al., 2006).

For each pair, the model estimated the expected values of the two cues, Q_A and Q_B, on the basis of individual sequences of choices and outcomes. Values were set at zero before learning and, after every trial $t > 0$, the value of the chosen cue (say A) was updated according to the Rescorla–Wagner rule: $Q_A(t+1) = Q_A(t) + \alpha * \delta(t)$. In the equation, $\delta(t)$ was the reward prediction error, calculated as $\delta(t) = R(t) - Q_A(t)$, and $R(t)$ was the reinforcement magnitude associated to the outcome of choosing cue A at trial t. Reinforcement magnitude was zero for "nothing" outcomes and adjusted as a free parameter (see below), positive for gains and negative for losses. Given the Q-values, the associated probability (or likelihood) of selecting each option was estimated by implementing the softmax rule, which is, for choosing A: $P_A(t) = exp(Q_A(t)/\beta)/(exp(Q_A(t)/\beta) + exp(Q_B(t)/\beta))$. Free parameters were individually adjusted to maximize the likelihood of observed choices, separately for the gain and the loss conditions. The search space was [0:0.1:1] for the learning rate (α), [0:0.1:1] for the choice randomness (β), and [0:0.1:2] for the reinforcement magnitude (with a positive sign for reward and negative for punishments). To test the effect of brain damage, we performed an ANOVA with group as the main factor, followed by post hoc between-group comparisons performed separately on the different free parameters using two-sample t tests.

REFERENCES

Atallah, H.E., Lopez-Paniagua, D., Rudy, J.W., and O'Reilly, R.C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. Nat. Neurosci. 10, 126–131.

Bar-Gad, I., and Bergman, H. (2001). Stepping out of the box: information processing in the neural networks of the basal ganglia. Curr. Opin. Neurobiol. 11, 689–695.

Blair, K., Marsh, A.A., Morton, J., Vythilingam, M., Jones, M., Mondillo, K., Pine, D.C., Drevets, W.C., and Blair, J.R. (2006). Choosing the lesser of two evils, the better of two goods: specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. J. Neurosci. 26, 11379–11386.

Bódi, N., Kéri, S., Nagy, H., Moustafa, A., Myers, C.E., Daw, N., Dibó, G., Takáts, A., Bereczki, D., and Gluck, M.A. (2009). Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain 132, 2385–2395.

Büchel, C., Morris, J., Dolan, R.J., and Friston, K.J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20, 947–957.

Craig, A.D. (2003). Interoception: the sense of the physiological condition of the body. Curr. Opin. Neurobiol. 13, 500–505.

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

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron 69, 1204–1215.

De Martino, B., Camerer, C.F., and Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. Proc. Natl. Acad. Sci. USA 107, 3788–3792.

Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C., and Fiez, J.A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. J. Neurophysiol. 84, 3072–3077.

Douaud, G., Gaura, V., Ribeiro, M.J., Lethimonnier, F., Maroy, R., Verny, C., Krystkowiak, P., Damier, P., Bachoud-Levi, A.C., Hantraye, P., and Remy, P. (2006). Distribution of grey matter atrophy in Huntington's disease patients: a combined ROI-based and voxel-based morphometric study. Neuroimage 32, 1562–1575.

Doya, K. (2002). Metalearning and neuromodulation. Neural Netw. 15, 495–506.

Draganski, B., Kherif, F., Klöppel, S., Cook, P.A., Alexander, D.C., Parker, G.J.M., Deichmann, R., Ashburner, J., and Frackowiak, R.S.J. (2008). Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. J. Neurosci. 28, 7143–7152.

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

Galvan, A., Hare, T.A., Davidson, M., Spicer, J., Glover, G., and Casey, B.J. (2005). The role of ventral frontostriatal circuitry in reward-based learning in humans. J. Neurosci. 25, 8650–8656.

Gottfried, J.A., and Dolan, R.J. (2004). Human orbitofrontal cortex mediates extinction learning while accessing conditioned representations of value. Nat. Neurosci. 7, 1144–1152.

Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. J. Chem. Neuroanat. 26, 317–330.

Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., and Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J. Neurosci. 28, 5623–5630.

Jensen, J., McIntosh, A.R., Crawley, A.P., Mikulis, D.J., Remington, G., and Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. Neuron 40, 1251–1257.

Jones, C.L., Ward, J., and Critchley, H.D. (2010). The neuropsychological impact of insular cortex lesions. J. Neurol. Neurosurg. Psychiatry 81, 611–618.

Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and Bergman, H. (2008). Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. J. Neurosci. 28, 11673–11684.

Kahnt, T., Park, S.Q., Cohen, M.X., Beck, A., Heinz, A., and Wrase, J. (2009). Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. J. Cogn. Neurosci. 21, 1332–1345.

Kim, H., Shimojo, S., and O'Doherty, J.P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. PLoS Biol. 4, e233.

Kim, H., Shimojo, S., and O'Doherty, J.P. (2011). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. Cereb. Cortex 21, 769–776.

Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., Ott, U., Burkart, J., and Vaitl, D. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: an event-related fMRI study. Neuroimage 20, 1086–1095.

Kuhnen, C.M., and Knutson, B. (2005). The neural basis of financial risk taking. Neuron 47, 763–770.

McCabe, C., Mishor, Z., Cowen, P.J., and Harmer, C.J. (2010). Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. Biol. Psychiatry 67, 439–445.

Miyazaki, K., Miyazaki, K.W., and Doya, K. (2011). Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. J. Neurosci. 31, 469–479.

Morrison, S.E., and Salzman, C.D. (2009). The convergence of information about rewarding and aversive stimuli in single neurons. J. Neurosci. 29, 11471–11483.

Naqvi, N.H., and Bechara, A. (2009). The hidden island of addiction: the insula. Trends Neurosci. 32, 56–67.

Nitschke, J.B., Dixon, G.E., Sarinopoulos, I., Short, S.J., Cohen, J.D., Smith, E.E., Kosslyn, S.M., Rose, R.M., and Davidson, R.J. (2006). Altering expectancy dampens neural response to aversive taste in primary taste cortex. Nat. Neurosci. 9, 435–442.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 4, 95–102.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304, 452–454.

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

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

Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., and Pessiglione, M. (2009b). Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. Proc. Natl. Acad. Sci. USA 106, 19179–19184.

Palminteri, S., Clair, A.-H., Mallet, L., and Pessiglione, M. (2012). Similar improvement of reward and punishment learning by serotonin reuptake inhibitors in obsessive-compulsive disorder. Biol. Psychiatry 72, 244–250.

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

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R.J., and Frith, C.D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. Neuron 59, 561–567.

Petrovic, P., Pleger, B., Seymour, B., Klöppel, S., De Martino, B., Critchley, H., and Dolan, R.J. (2008). Blocking central opiate function modulates hedonic impact and anterior cingulate response to rewards and losses. J. Neurosci. 28, 10509–10516.

Postuma, R.B., and Dagher, A. (2006). Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. Cereb. Cortex 16, 1508–1521.

Ravel, S., Legallet, E., and Apicella, P. (2003). Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. J. Neurosci. 23, 8489–8497.

Rutledge, R.B., Dean, M., Caplin, A., and Glimcher, P.W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. J. Neurosci. 30, 13525–13536.

Samanez-Larkin, G.R., Hollon, N.G., Carstensen, L.L., and Knutson, B. (2008). Individual differences in insular sensitivity during loss anticipation predict avoidance learning. Psychol. Sci. 19, 320–323.

Schlund, M.W., and Cataldo, M.F. (2010). Amygdala involvement in human avoidance, escape and approach behavior. Neuroimage 53, 769–776.

Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., Friston, K.J., and Frackowiak, R.S. (2004). Temporal difference models describe higher-order learning in humans. Nature *429*, 664–667.

Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. J. Neurosci. *27*, 4826–4831.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (Cambridge, MA: The MIT Press).

Tabrizi, S.J., Langbehn, D.R., Leavitt, B.R., Roos, R.A., Durr, A., Craufurd, D., Kennard, C., Hicks, S.L., Fox, N.C., Scahill, R.I., et al.; TRACK-HD investiga-tors. (2009). Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. Lancet Neurol. *8*, 791–801.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. Science *315*, 515–518.

Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. J. Risk Uncertain. *5*, 297–323.

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D.F., and Büchel, C. (2006). Dissociable systems for gain- and loss-related value predic-tions and errors of prediction in the human brain. J. Neurosci. *26*, 9530–9537.