

Beyond Reversal: A Critical Role for Human Orbitofrontal Cortex in Flexible Learning from Probabilistic Feedback

Ami Tsuchida,¹ Bradley B. Doll,² and Lesley K. Fellows¹

¹Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada H3A 2B4, and ²Department of Cognitive, Linguistic, and Psychological Sciences, Brown Institute for Brain Science, Brown University, Providence, Rhode Island 02912

Damage to the orbitofrontal cortex (OFC) has been linked to impaired reinforcement processing and maladaptive behavior in changing environments across species. Flexible stimulus–outcome learning, canonically captured by reversal learning tasks, has been shown to rely critically on OFC in rats, monkeys, and humans. However, the precise role of OFC in this learning remains unclear. Furthermore, whether other frontal regions also contribute has not been definitively established, particularly in humans. In the present study, a reversal learning task with probabilistic feedback was administered to 39 patients with focal lesions affecting various sectors of the frontal lobes and to 51 healthy, demographically matched control subjects. Standard groupwise comparisons were supplemented with voxel-based lesion–symptom mapping to identify regions within the frontal lobes critical for task performance. Learning in this dynamic stimulus–reinforcement environment was considered both in terms of overall performance and at the trial-by-trial level. In this challenging, probabilistic context, OFC damage disrupted both initial and reversal learning. Trial-by-trial performance patterns suggest that OFC plays a critical role in interpreting feedback from a particular trial within the broader context of the outcome history across trials rather than in simply suppressing preexisting stimulus–outcome associations. The findings show that OFC, and not other prefrontal regions, plays a necessary role in flexible stimulus–reinforcement learning in humans.

Introduction

The optimal adjustment of behavior in dynamic environments is fundamental to adaptive behavior. Evidence across species points to the critical involvement of orbitofrontal cortex (OFC) in the flexible control of behavior in changing environments (Rolls, 2004; Schoenbaum et al., 2009). In rodents and nonhuman primates, lesions of OFC impair the ability to adjust behavior in response to changing stimulus–reward contingencies, as in reversal learning tasks (Iversen and Mishkin, 1970; Jones and Mishkin, 1972; Dias et al., 1996; Schoenbaum et al., 2003). In these tasks, animals learn to associate a stimulus with either reward or non-reward. After the stimulus–reinforcement associations are established, reinforcement contingencies are reversed, and the animal must overcome the previously established association to choose adaptively. Nonhuman primates with lesions to OFC, but not those with lesions to dorsolateral prefrontal cortex (DLPFC), are slow to learn the new contingencies, making perseverative responses to previously rewarded stimuli (Dias et al., 1996). A sim-

ilar pattern of impairment in reversal learning from deterministic feedback has been demonstrated in human patients with focal OFC damage (Fellows and Farah, 2003).

In contrast to this lesion work, functional imaging studies of healthy subjects have focused on reversal learning under the more challenging (and possibly more ecologically relevant) conditions of probabilistic feedback (Cools et al., 2002; O'Doherty et al., 2001, 2003; Remijne et al., 2005; Budhani et al., 2007). Although some of these studies do report feedback-related activity in OFC regions, other areas of PFC have also been implicated: for example, activations related to response reversal after contingency change have been reported in DLPFC, dorsal anterior cingulate (dACC), and anterior prefrontal regions (O'Doherty et al., 2003; Remijne et al., 2005). Importantly, probabilistic environments may make additional demands on learning and evaluation systems (Rudebeck et al., 2008; Walton et al., 2010), including the development of more abstract strategies (Genovesio et al., 2005), and the extent to which OFC and other prefrontal regions are critical for adaptive learning under such conditions is not clear. Evidence from existing neuropsychological studies of human subjects with focal frontal lobe lesions does point to the importance of OFC (Berlin et al., 2004; Hornak et al., 2004), at least in the reversal phase of such learning, but DLPFC has also been implicated (Hornak et al., 2004).

To investigate the specific contributions of OFC and other prefrontal regions to adaptive learning in a dynamic environment, we administered a reversal learning task with probabilistic feedback to 39 patients with focal frontal lobe damage and 51 healthy subjects. Standard region-of-interest analyses focusing on OFC were supplemented by voxelwise lesion-

Received April 8, 2010; revised Aug. 4, 2010; accepted Aug. 13, 2010.

The work was supported by Canadian Institutes of Health Research (CIHR) Grant MOP-77583, a Le Fonds de la Recherche en Santé de Québec operating grant, and a CIHR Clinician–Scientist Award (all to L.K.F.) and a Frederick Banting and Charles Best Canada Graduate Scholarships Doctoral Award (A.T.). We thank Michael Frank for thoughtful comments on the trial-by-trial analysis. Marianna Stark and Arlene Berg were instrumental in recruiting and screening patients through the patient databases at the University of Pennsylvania and McGill University. Institutional support of those two resources is gratefully acknowledged, as is the ongoing commitment of patients, families, and referring clinicians.

Correspondence should be addressed to Ami Tsuchida, Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, QB, Canada H3A 2B4. E-mail: ami.tsuchida@mail.mcgill.ca.

DOI:10.1523/JNEUROSCI.1958-10.2010

Copyright © 2010 the authors 0270-6474/10/3016868-08\$15.00/0

function mapping to identify prefrontal regions critical for task performance. In addition, trial-by-trial performance was examined. The results confirmed an important role for OFC in adaptive reinforcement learning from probabilistic feedback, indicating that this region is critical for both initial and reversal learning under conditions of uncertainty. Human OFC, but not other prefrontal regions, appears to play a critical, general role in reinforcement learning in dynamic, probabilistic reinforcement contexts.

Materials and Methods

Subjects. Subjects with focal lesions involving the frontal lobes ($n = 39$) were recruited from research databases at the Center for Cognitive Neuroscience at the University of Pennsylvania and McGill University. They were eligible if they had a fixed lesion primarily affecting the frontal lobes. They were tested a minimum of 6 months after the injury (mean, 4.6 years after; range, 9 months to 16 years). Sixteen subjects were taking psychoactive medications. These were most commonly anticonvulsants or antidepressants.

Age- and education-matched healthy control (CTL) subjects ($n = 51$) were recruited through local advertisement in Montreal. They were free of neurological or psychiatric disease and were not taking any psychoactive drugs. They were excluded if they scored <28 of 30 on the minimal status examination (Folstein et al., 1975) or <26 of 30 on the Montreal Cognitive Assessment (Nasreddine et al., 2005). All subjects provided written, informed consent in accordance with the Declaration of Helsinki and were paid a nominal fee for their time. The study protocol was approved by the Institutional Review Boards of both participating centers.

Probabilistic reversal learning task. The ability to learn stimulus–reinforcement associations and to reverse them based on probabilistic feedback was measured using a simple computerized card game similar in format to a deterministic reversal learning task used in a previous study (Fellows and Farah, 2003). Subjects were asked to pick a card from either a blue or an orange deck by touching it on a touch-screen-equipped computer. After each choice, feedback of \$50 (play money) win or \$50 (play money) loss was provided. A score bar was also displayed, indicating the cumulative amount of play money won by the subject, updated after each trial. One deck was associated with a win on six of every seven trials (i.e., 86%) and a loss on one in seven trials (14%). The other deck had the opposite pattern. We refer to the deck associated more often with wins as the “advantageous” deck and the other as the “disadvantageous” deck. For the purposes of the trial-by-trial analysis, we defined feedback as congruent if it was consistent with the overall value of the chosen deck (i.e., a win resulting from choosing the currently advantageous deck or a loss after the choice of the currently disadvantageous deck). In contrast, incongruent feedback was defined as an outcome inconsistent with the overall value of the chosen deck (supplemental Table S1, available at www.jneurosci.org as supplemental material).

Once the subject reached criterion by choosing the advantageous deck 13 of any 14 consecutive trials, the overall stimulus–reward contingencies reversed (i.e., the advantageous deck became disadvantageous and vice versa). If the learning criterion was met again, the contingencies reversed again for a total of 70 trials, allowing for up to five reversals. If the subject failed to reach the initial learning criterion within 50 trials, the task was terminated. Although subjects learned these contingencies by trial and error, they were given instructions that indicated the probabilistic nature of the feedback, as follows: “You are going to play a simple card game for play money. Your goal is to win as much money as you can and to lose as little as possible. On each turn, you will pick a card from one of two decks. Neither deck is an ‘all-winning’ deck, but one of them is better than the other. You will do well if you STICK WITH the better deck, but be careful because the better deck may change.”

Lesion analysis. Individual lesions were traced from the most recent clinical computed tomography or magnetic resonance imaging onto the standard Montreal Neurological Institute (MNI) brain using MRIcro software (Rorden and Brett, 2000) (freely available at www.mricro.com) by a neurologist experienced in imaging analysis and blind to task per-

formance. The same software was used to estimate lesion volumes and to generate lesion overlap images. The subjects with frontal lobe damage were first divided a priori into those with a lesion affecting OFC (OFC group) and those who had damage elsewhere. The latter group was further divided into medial (MF) and lateral (LF) PFC groups, to test for differences in performance between these two non-overlapping groups. Lesions affecting OFC were attributable to aneurysm rupture in six cases, tumor resection in three, and ischemic or hemorrhagic stroke in three. MF lesions were attributable to tumor resection in 9 of 13 cases, with the rest attributable to ischemic stroke. Lesions in LF group were caused by ischemic or hemorrhagic stroke in 12 cases and tumor resection in two.

Region-of-interest statistical analysis. The effect of group membership on overall performance was assessed by submitting total scores to a between-subject ANOVA. We used the cumulative score (the total amount of play money won) reached by trial 83, which was the minimum number of trials completed by all those who passed the initial learning stage, as the primary measure of overall performance.

To determine whether the difference in overall performance was specifically attributable to difficulty during reversal or to learning from probabilistic feedback more generally, the number of errors made in the initial and reversal learning phases was submitted to repeated-measures ANOVA, with learning phase as within-subject factor. [Errors were choices from the currently disadvantageous deck, regardless of the feedback received by the subject on that particular trial (supplemental Table S1, available at www.jneurosci.org as supplemental material)]. Because subjects achieved a different number of reversals, depending on how well they performed, the average number of errors made per reversal was the measure of interest. Previous work with a deterministic version of the task found that most of errors were made immediately after reversal, with some subjects adapting more slowly to the contingency changes than others. To specifically test for group differences in such perseverative responses after reversal in this probabilistic context, the number of consecutive losses required for the subject to choose the alternate deck after each reversal was determined and submitted to repeated-measures ANOVA.

In addition to these more global measures, differential patterns of trial-by-trial performance were examined by comparing group differences in win–shift or lose–shift behavior, an approach taken in recent functional magnetic resonance imaging studies (Jocham et al., 2009a; Kahnt et al., 2009). In this probabilistic reinforcement environment, win–shift or lose–shift behavior could be adaptive or maladaptive, depending on the context. To capture these distinctions, we distinguished between shift responses after congruent feedback and those after incongruent feedback. Although there is no way for a subject to tell the difference between congruent and incongruent feedback in any given trial in isolation, this can be inferred from the recent reinforcement history of the chosen deck. Indeed, optimal performance of this task requires interpreting feedback within this broader context. This trial-by-trial measure is an index of how well participants were able to do so.

The frequency of win–shift behavior was calculated by counting the number of response shifts immediately after congruent positive feedback and dividing it by the total number of trials after congruent positive feedback. This measure was then submitted to between-subject ANOVA. There were too few trials with incongruent positive feedback to allow these to be analyzed. (Subjects have to choose the currently disadvantageous deck often to experience any incongruent positive feedback.) For the analysis of lose–shift behavior, the frequency of adaptive shifts (i.e., shifting in response to congruent negative feedback) and maladaptive shifts (i.e., shifting in response to incongruent negative feedback) was calculated by counting the number of response shifts after congruent or incongruent negative feedback and dividing it by the total number of trials after each form of negative feedback. These were then submitted to repeated-measures ANOVA, with type of negative feedback as a within-subject factor.

When analyses revealed significant group effects, *post hoc* pairwise comparisons were made with the Tukey–Kramer method. When significant interactions were detected, simple main effects tests were performed, followed by Tukey–Kramer pairwise comparisons, when applicable.

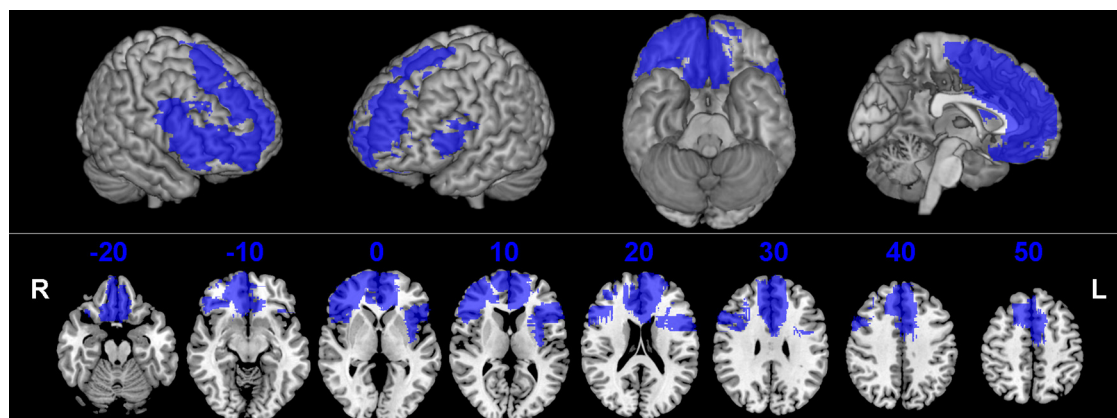


Figure 1. Map showing the voxels (blue) where there is sufficient statistical power to detect an effect in this group of patients, overlaid on the MNI brain and shown in three-dimensional views and in axial slices, with numbers indicating the z-coordinates (MNI) of each slice. L, Left; R, right.

Table 1. Demographic information for CTL and FP groups and for anatomically defined frontal subgroups (OFC, LF, and MF) (mean \pm SD)

Group	Age (years)	Education (years)	Sex (F/M)	BDI	ANART IQ	Lesion volume (cc)
CTL ($n = 48$)	56.4 \pm 12.4	15.1 \pm 2.9	18/30	5.0 \pm 4.9	123.1 \pm 7.7	
FP ($n = 36$)	56.9 \pm 11.6	14.1 \pm 3.5	15/21	12.7 \pm 9.9*	115.9 \pm 9.2*	38.3 \pm 42.5
OFC ($n = 11$)	60.5 \pm 9.2	13.1 \pm 3.9	2/9	14.0 \pm 12.8	112.7 \pm 10.5	50.2 \pm 70.0
LF ($n = 12$)	56.4 \pm 12.6	14.3 \pm 2.3	4/8	12.4 \pm 8.8	114.0 \pm 10.0	37.0 \pm 19.6
MF ($n = 13$)	54.5 \pm 12.5	14.8 \pm 4.0	9/4	11.8 \pm 8.6	119.9 \pm 6.4	29.4 \pm 24.7

Not all patients were able to complete the ANART. F, Female; M, male. * $p < 0.05$, significant differences based on a two-tailed t test for comparisons between CTL and FP or an ANOVA for comparison across the three frontal subgroups.

Table 2. Performance on neuropsychological screening tests for CTL and FP groups (mean \pm SD)

Group	Sentence comprehension accuracy	Fluency–F	Fluency–animal	Backward digit span	Backward Corsi span
CTL ($n = 48$)	0.98 \pm 0.07	14.3 \pm 4.9	21.5 \pm 5.1	5.0 \pm 1.4	4.7 \pm 1.1
FP ($n = 36$)	0.96 \pm 0.06	9.8 \pm 5.3*	15.8 \pm 4.9*	4.2 \pm 1.8*	4.0 \pm 1.8*
OFC ($n = 11$)	0.94 \pm 0.07	9.9 \pm 4.7	15.1 \pm 4.5	3.7 \pm 0.9	3.5 \pm 1.4
LF ($n = 12$)	0.96 \pm 0.07	9.3 \pm 5.7	14.8 \pm 4.9	4.0 \pm 1.3	4.5 \pm 0.9
MF ($n = 13$)	0.98 \pm 0.05	10.2 \pm 5.6	17.2 \pm 5.3	4.8 \pm 1.6	4.0 \pm 1.5

* $p < 0.05$, significant differences (t test or ANOVA).

Behavior-based statistical analysis. The region-of-interest approach may miss effects of damage to prefrontal regions outside OFC and does not identify the specific areas within OFC that may be responsible for any group effects. Voxel-based lesion–symptom mapping (VLSM) was therefore applied to test for lesion–behavior associations without a priori assumptions about the critical anatomical regions. In VLSM, the performance measure of interest is entered as a continuous measure, and statistical comparisons are made for each eligible voxel, comparing the performance of subjects with damage affecting a given voxel with that of subjects with damage outside of that voxel. The Brunner–Munzel (BM) test (Brunner and Munzel, 2000), a nonparametric method, was used to compare performance on a voxelwise basis, as implemented in NPM (version 29, October 2008) and MRICroN (version 15, October 2008) software (Rorden et al., 2007) (www.mricro.com/npm/ for NPM and www.mricro.com/mricron for MRICroN). Note that this version of the software appropriately corrects for small group size. Only voxels affected in at least three cases were included for this analysis, and a cluster extent threshold of $k = 50$ voxels was applied. Statistical maps thresholded at an uncorrected p value < 0.05 were generated to show the relative effects of lesions to various regions within PFC. Figure 1 shows the statistical power map, indicating the voxels where we had adequate power to detect effects at this uncorrected threshold. The map is analogous to the “effective coverage map” described by Rudrauf et al. (2008) for lesion mapping with binomial behavioral data but uses the Wilcoxon–Mann–Whitney probability to estimate power for continuous behavioral data (Gläscher et al., 2009).

Results

Three of 39 patients with frontal lobe damage and 3 of 51 healthy control participants failed to reach the initial learning

criterion ($\chi^2_{(1)} = 0.102$, $p > 0.05$). Supplemental Figure S1 (available at www.jneurosci.org as supplemental material) shows the lesion locations in these patients, which do not seem to follow any particular pattern. These subjects were excluded from additional analyses.

Table 1 provides demographic and background information for the remaining participants. Intelligence quotient (IQ) was estimated with the American version of National Adult Reading Test (ANART). Unpaired t tests revealed no significant difference between frontal patients (FP) and controls for age or education, as intended ($p > 0.05$). The frontal group as a whole had significantly lower estimated premorbid IQ and higher Beck Depression Inventory (BDI) scores compared with controls ($p < 0.05$), but there were no differences in these scores or in any of the demographic variables across the three a priori anatomically defined subgroups of frontal patients (ANOVA, all $p > 0.05$).

Cognitive abilities of no interest in this study, but which may be affected by frontal lobe damage, were measured in all patients by means of a brief screening neuropsychological evaluation. Tasks included sentence comprehension, letter and category fluency tests, backward digit span, and a modified, computerized version of the Corsi block tapping task. The results of these tests are summarized in Table 2. As expected, the frontal group as a whole performed significantly worse than control subjects in all of the tests ($p < 0.01$) except for sentence comprehension ($p > 0.05$). However, there were no differences in performance of any of

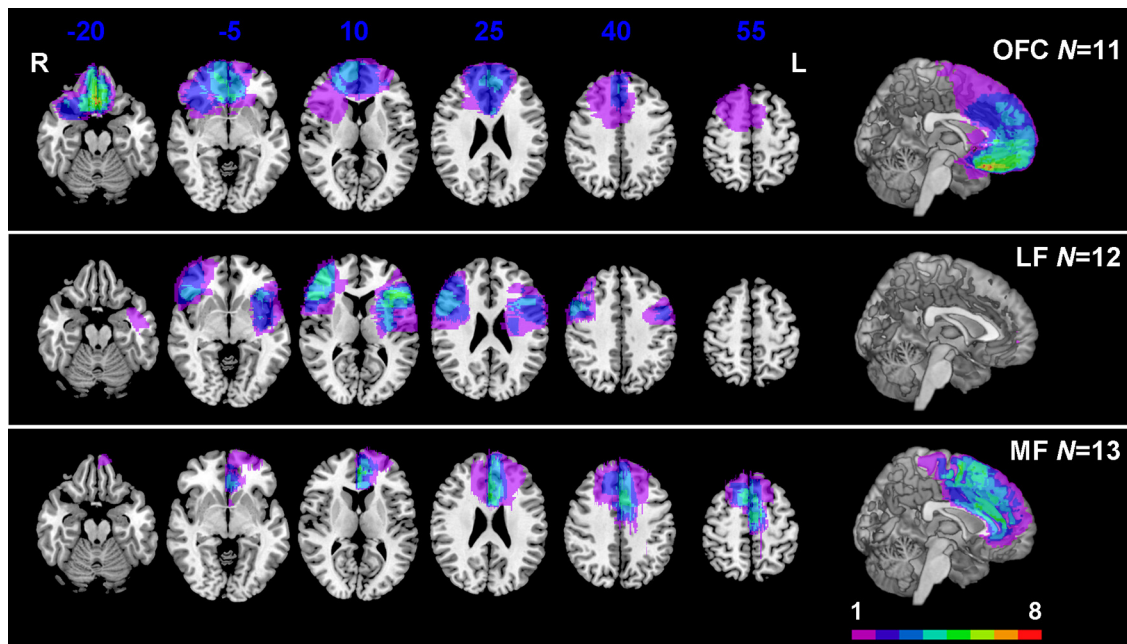


Figure 2. Representative axial slices and midsagittal views of the MNI brain, showing the degree of lesion overlap for subjects with damage affecting OFC (OFC group, $n = 11$; top row) and those without damage to this region, divided into those with lesions involving lateral prefrontal cortex (LF group, $n = 12$; second row) or the medial surface of the prefrontal cortex (MF group, $n = 13$; bottom row). Colors indicate the degree of overlap across subjects, as shown in the legend. L, Left; R, right.

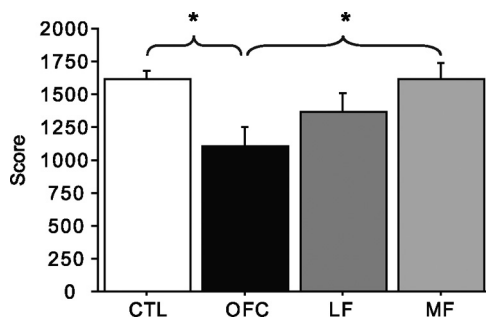


Figure 3. Total probabilistic reversal learning task scores for each group. Error bars indicate SEM. * $p < 0.05$, significant differences for pairwise comparisons (Tukey–Kramer test).

these tests between the three subgroups of frontal patients (ANOVA or Kruskal–Wallis tests, all $p > 0.05$).

Overall performance on the reversal learning task was compared across anatomical subgroups and healthy control subjects (OFC group, $n = 11$; LF group, $n = 12$; MF group, $n = 13$; CTL group, $n = 48$). Figure 2 shows the lesion locations, as overlap images, for the three anatomically defined groups. A one-factor ANOVA with group as the between-subject factor revealed a significant effect of group for the total score ($F_{(3,80)} = 4.44$, $p = 0.006$), with the *post hoc* comparisons identifying a significantly lower score in the OFC group relative to both the CTL and MF groups (Tukey–Kramer test, $p < 0.05$) (Fig. 3). This general pattern was also true for other measures of performance, including total number of reversals and total number of choices from the disadvantageous deck, both of which were tightly correlated with the total score (Table 3). Subjects with damage to OFC thus performed poorly on the task overall. This replicates the previous finding with a deterministic version of this task (Fellows and Farah, 2003). Although three subjects in the OFC group and five in the LF or MF groups participated in both studies, significant impairment in task performance in those with OFC damage was present even when

the analysis was restricted to those who had not participated in the study reported by Fellows and Farah (2003), constituting replication of this finding in an independent sample.

Although the region-of-interest analysis argues against any substantial contribution of MF or LF areas to task performance, potential effects may be masked or diluted if putatively critical regions are poorly captured by the regional divisions imposed in the primary analysis. To explore whether other prefrontal regions contribute to task performance and to better delineate the region within OFC that was responsible for the observed group effect, a VLSM analysis was undertaken. The total score reached by the patients with focal frontal lobe damage was entered as the dependent variable to identify voxels that, when lesioned, were associated with low scores on the reversal learning task.

VLSM methods are relatively new, and there is ongoing debate about the most suitable implementation of the statistics (Rorden et al., 2009). As a preliminary step, we applied the nonparametric Brunner–Munzel test at each voxel affected in at least three participants. The statistical map thus generated is shown in Figure 4, thresholded at $p < 0.05$, uncorrected. No voxels survived correction for multiple comparisons using conventional false discovery rate or permutation thresholding. In principle, the uncorrected map should be interpreted with caution because of the risk of false-positive findings. However, interpreted in the context of the primary region-of-interest results, which demonstrated a significant effect of OFC damage, this map provides useful insights. The VLSM results indicate that poor performance in the predefined OFC group is driven mainly by damage to voxels in bilateral posteromedial OFC. There were other, smaller clusters of voxels associated with poor performance to a weaker degree, as listed in Table 4.

The VLSM map must also be interpreted within the context of what it cannot show, i.e., associations cannot be detected in those regions of prefrontal cortex that were not adequately sampled in the patient group we studied (Kimberg et al., 2007). Figure 1 shows the statistical power map indicating where we had ade-

Table 3. Summary of alternative behavioral measures in the reversal learning task (mean \pm SD)

	CTL	OFC	LF	MF	Group comparison	Post hoc comparison
Total number of reversals	4.0 \pm 1.0	2.9 \pm 0.8	3.4 \pm 1.4	3.8 \pm 1.4	$p = 0.027$	OFC < CTL
Total number of errors	16.5 \pm 7.0	23.9 \pm 5.6	17.8 \pm 5.5	17.5 \pm 6.7	$p = 0.015$	OFC > CTL
Errors in initial learning	2.3 \pm 2.2	4.5 \pm 2.8	3.1 \pm 2.1	2.7 \pm 1.7	$p = 0.037$	OFC > CTL
Errors in first reversal	5.9 \pm 6.3	9.7 \pm 7.1	8.1 \pm 7.9	6.4 \pm 7.1	$p = 0.060$	OFC > CTL

Group comparison were made with ANOVA, and the *post hoc* comparison column indicates when significant differences between groups were detected with the Tukey–Kramer test ($p < 0.05$).

quate power to detect effects at the uncorrected threshold applied in Figure 4. At this threshold, sufficient power was available to detect associations with most regions within PFC, including the OFC areas in which we found significant effects as well as broad regions of medial PFC and right lateral PFC. The only regions where we did not have adequate power to detect effects at this threshold were sections of the superior and middle frontal gyri bilaterally and left lateral OFC.

Having established that OFC damage is selectively associated with poor performance on the task as a whole and having clarified which voxels within OFC were driving this effect, we next asked whether this deficit was specific to the reversal learning phase of the task, as has been suggested by previous work (Rolls et al., 1994; Fellows and Farah, 2003; Hornak et al., 2004). In the probabilistic reinforcement context studied here, subjects with OFC damage performed poorly overall, with their errors similarly distributed over initial and reversal phases of the task. When errors in the initial and reversal phases were submitted to repeated-measures ANOVA, with group as between-subject factor and learning phase as within-subject factor, a significant effect of learning phase ($F_{(1,80)} = 21.65$, $p < 0.0001$) and a significant effect of group ($F_{(3,80)} = 3.00$, $p = 0.04$) were present, but there was no significant group \times learning phase interaction ($F_{(3,80)} = 0.58$, $p = 0.63$) (Fig. 5A). The OFC group made more errors than control subjects overall (Tukey–Kramer test, $p < 0.05$), regardless of learning phase.

Furthermore, and again in contrast to their performance in fully determined reinforcement environments, those with OFC damage did not exhibit exaggerated perseverative behavior after reversal. The number of consecutive negative outcomes that occurred before the subject switched to the other option in the initial learning and reversal learning phases did not differ across groups, nor was there a significant group \times learning phase interaction ($F_{(3,80)} = 0.49$, 0.49 and $p = 0.69$, 0.69 , respectively) (Fig. 5B). The number of choices followed by negative feedback required for subjects in the OFC group to try another deck was essentially identical to that of the CTL group in both initial and reversal phases.

If the deficits in probabilistic reinforcement learning in OFC patients are not the manifestation of simple perseveration despite negative feedback, then what is the trial-by-trial basis of this global impairment? This question was addressed by examining the trialwise performance of each subject to compare the effects of positive or negative feedback on the next choice, across groups. First, to investigate the effect of a positive outcome on the pre-

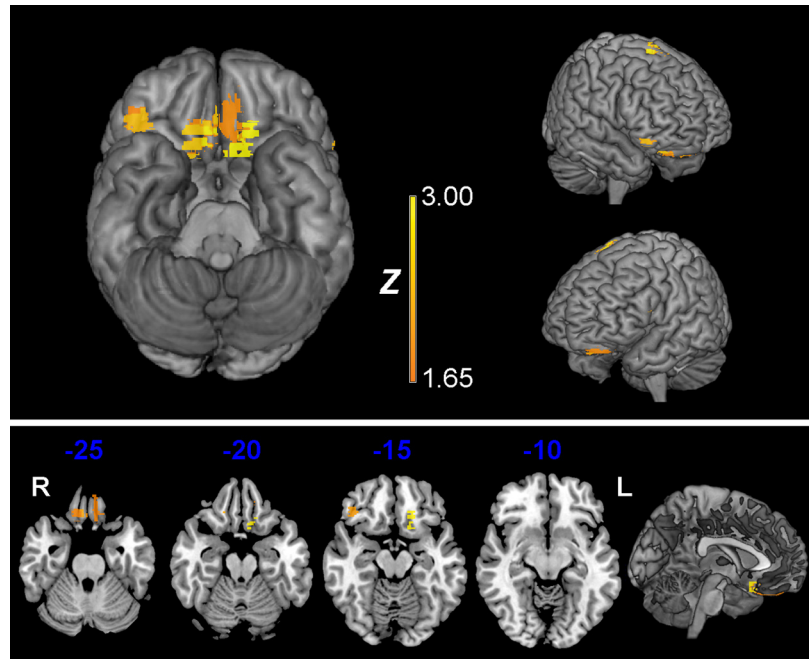


Figure 4. VLSM statistical map computed for total score shown on three-dimensional views of the MNI brain (top), with representative axial slices (bottom). The color scale indicates BM test Z-score, thresholded at $p < 0.05$, uncorrected.

ceding trial, the frequency of win–shift behavior was submitted to between-subjects ANOVA. This revealed a significant effect of group ($F_{(3,80)} = 2.94$, $p = 0.04$; Kruskal–Wallis test, $p = 0.004$), with *post hoc* comparisons showing increased win–shift behavior in the OFC group compared with the CTL group (Tukey–Kramer test, $p < 0.05$) (Fig. 6A). Thus, although the CTL group tended to be adaptively influenced by a positive outcome, choosing the same deck again after winning money on the previous trial, the OFC group was more likely to switch their choice after winning.

For the analysis of behavior after negative feedback, we distinguished between congruent and incongruent (i.e., misleading) negative feedback to determine whether there were any group differences in how well negative feedback was interpreted in context, in addition to any overall differences in sensitivity to negative feedback. A repeated-measures ANOVA with the type of negative feedback (i.e., congruent and incongruent) as a within-subject factor did not reveal any group difference in overall lose–shift behavior ($F_{(3,80)} = 1.18$, $p = 0.32$) but indicated a tendency for a group by feedback type interaction ($F_{(3,80)} = 2.21$, $p = 0.09$) (Fig. 6B). Closer examination using simple main effects tests revealed that, although CTL, MF, and LF groups all showed a strong effect of negative feedback type on shifting behavior on the next trial ($F_{(1,80)} > 11.38$, all p values < 0.001), there was no such effect in the OFC group ($F_{(1,80)} = 0.31$, $p = 0.58$). In other words, although CTL and other patient groups were able to adaptively adjust their shifting behavior depending on the context (shifting more often after congruent than incongruent negative feedback),

Table 4. Coordinates of the regions associated with impaired performance in the VLSM analysis, in MNI space

Region	Hemisphere	x	y	z	BM Z maximum	n voxels
Gyrus rectus	Left	−6	21	−24	2.85	1342
Gyrus rectus	Right	7	23	−24	2.33	168
Superior frontal gyrus, orbital part	Left	−14	15	−18	2.85	509
Inferior frontal gyrus, orbital part	Right	39	37	−15	2.01	408
Inferior frontal gyrus, orbital part, white matter	Left	−20	34	1	2.33	60
Supplementary motor area	Right	9	6	61	2.44	147
Supplementary motor area	Right	4	15	70	2.44	53

Region labels are taken from the automated anatomical labeling template (Tzourio-Mazoyer et al., 2002). The MNI coordinates indicate the center of mass for each significant cluster. The maximum BM Z statistics obtained for each cluster are also shown. Z scores greater than 2.3 are significant at $p < 0.01$ (BM test, uncorrected) and greater than 1.6 at $p < 0.05$ (uncorrected).

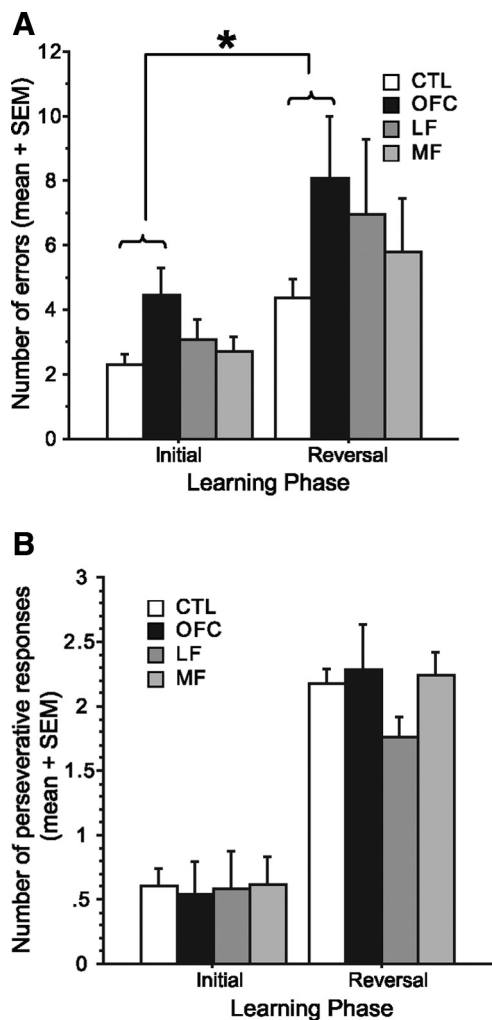


Figure 5. *A*, The number of errors during initial and reversal learning phase in each group. *B*, The number of consecutive responses to punished choices before switching to the other option in the initial learning phase of the task and after reversals. Error bars indicate SEM. * $p < 0.05$, significant differences for pairwise comparisons (Tukey–Kramer test).

the OFC group was less able to interpret the meaning of negative feedback in the context of the broader reinforcement history.

Discussion

The present study found that damage to OFC, but not to other areas within PFC, resulted in impaired performance on a probabilistic reversal learning task. This is consistent with two previous neuropsychological studies that reported impaired performance of a complex, gambling-like reversal learning task after OFC damage (Berlin et al., 2004; Hornak et al., 2004). The large cohort studied here allowed VLSM analysis to be applied to much of the

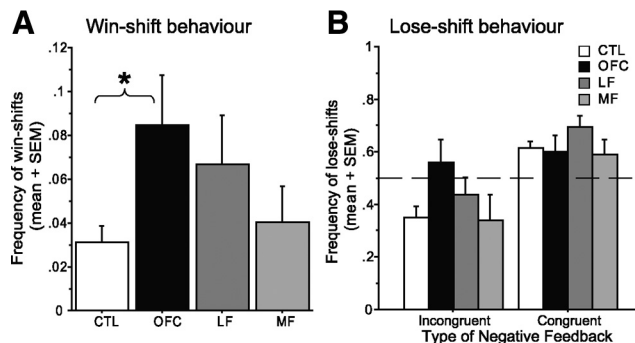


Figure 6. Trialwise performance analyses. *A*, The frequency of win–shift behavior in each group. *B*, The frequency of lose–shift behavior after congruent or incongruent negative feedback (for details, see Results). Error bars indicate SEM. * $p < 0.05$, significant differences for pairwise comparisons (Tukey–Kramer test).

PFC. We were able to confirm a regionally specific contribution of OFC, with the effect mainly driven by voxels in bilateral posteromedial OFC and to a lesser extent right lateral OFC, and to reject a critical role for other regions within PFC, notably including dACC (at least of the same effect size and within the anatomical constraints of our sample) in flexible reinforcement learning in a probabilistic environment.

Contrary to the performance of such patients on a simple, deterministic reversal learning task (Fellows and Farah, 2003), the impairment of OFC patients was not restricted to the reversal phase of the task. Such patients made more errors even during initial learning in this more challenging, probabilistic environment. This is consistent with a general role for OFC in using information about outcome expectancy to adjust behavior: in this view, reversal of the stimulus–reward contingency is simply one instance of a general requirement for behavioral adjustment based on expectancy violation (Murray and Izquierdo, 2007) or the demand for precise linkage between a stimulus and a particular instance of reinforcement (Walton et al., 2010). In probabilistic contexts, accurate outcome expectancy requires integration of previous feedback history, titrated to the particular reinforcement structure of the task. Previous work supports a more general deficit in probabilistic learning after ventral frontal damage. One study using a simple probabilistic reversal learning task in patients with damage likely involving the ventral frontal lobe reported a tendency for increased errors during the acquisition phase (Swainson et al., 2000). A recent study investigating the effect of OFC damage on the ability to learn from positive and negative feedback used a challenging probabilistic learning task without reversal (Wheeler and Fellows, 2008). In that study, patients with OFC damage were poor at probabilistic learning, taking longer to reach criterion and never learning as well as healthy subjects or patients with frontal lobe damage sparing OFC, even

when reversal was not required. It is of interest to note that, in other studies using probabilistic reversal learning tasks, impaired performance not restricted to the reversal phase has also been observed in patients with amygdala damage (Hampton et al., 2007), patients with major depressive disorder (Murphy et al., 2003), and healthy subjects administered a central serotonin reuptake inhibitor (Chamberlain et al., 2006). The deficits reported in the latter two studies have been interpreted as reflecting altered OFC function, and our results support this claim.

One interpretation of the deficit in reversal learning after OFC lesions, whether in animals or humans, is that it is attributable to the loss of inhibitory control over responses to reward-related stimuli, resulting in perseverative responses to previously rewarded stimuli (Jones and Mishkin, 1972; Rolls et al., 1994; Dias et al., 1996; Chudasama et al., 2003). This is difficult to reconcile with the more general impairments in both initial and reversal learning reported here. Furthermore, despite a clear deficit in overall task performance, OFC patients in our study did not make more perseverative responses after reversal relative to control subjects or other frontal patient groups, in the sense that they did not keep choosing the previously rewarded stimulus longer than other groups before trying the previously nonrewarded stimulus. This pattern agrees with a recent report demonstrating that monkeys with restricted lesions to OFC exhibited a more pronounced deficit after they correctly chose the newly rewarded stimulus after reversal, not before (Rudebeck and Murray, 2008). The patterns of impairment in previous studies of OFC patients in probabilistic contexts also suggest that they make errors even after correctly choosing newly rewarded stimuli (Berlin et al., 2004; Hornak et al., 2004). Together, these data suggest that OFC is critical for reversal learning not because of a role in response inhibition or response shifting. Instead, these findings support the view that OFC is crucial in flexibly guiding behavior when outcome expectancies are violated (Murray et al., 2007; Schoenbaum et al., 2009), perhaps by playing a role in tightly linking a given trial to a given reinforcement outcome (Walton et al., 2010).

The analysis of trial-by-trial performance revealed an elevated frequency of win–shift trials in the OFC group. The inability of this group to sustain the correct choice after positive feedback is consistent with previous studies in OFC patients reporting “reward insensitivity” (Berlin et al., 2004; Hornak et al., 2004). Similar impairment has also been reported in monkeys with OFC lesions performing an object reversal learning task (Rudebeck and Murray, 2008). Although “punishment insensitivity” was also reported in previous studies of OFC patients (Berlin et al., 2004; Hornak et al., 2004), sensitivity to negative feedback can be adaptive or maladaptive in the context of probabilistic reversal learning. Our study clearly demonstrated that, although overall lose–shift behavior did not differ between groups, only the OFC group failed to adaptively tailor shifting behavior to congruent compared with incongruent negative feedback. This pattern may seem at odds with a previous lesion study demonstrating that OFC is critical for learning from negative, but not positive, feedback (Wheeler and Fellows, 2008). However, that study measured the ability to integrate cumulative positive or negative feedback experience over hundreds of trials, whereas what we measured here was the response to immediate feedback on individual trials, in a much briefer task. These two studies together argue that these two facets of feedback-driven learning are distinct.

This study does not directly address whether different areas within OFC play distinct roles in processing positive and negative feedback. Even with VLSM, we lack the spatial resolution likely to be needed to answer that question, and adaptive reversal learning

in this task asymmetrically hinges on interpreting negative feedback in context. Additional work is still needed to determine whether it is the valence of the feedback or the difficulty of interpreting it in context that drives the effects observed after OFC damage.

Our finding that the MF group, with lesions centered around dACC (Fig. 2), performed very similarly to the healthy control group was surprising, given the accumulating evidence for the importance of dACC region in performance monitoring and various reward-related effects on behavior (Ridderinkhof et al., 2004; Williams et al., 2004; Rushworth et al., 2007a). Based on functional imaging studies in humans and lesion work in monkeys, it has been suggested that this region is particularly important for choices in stochastic environments in which integration of outcome history is required for behavioral adjustment (Kennerly et al., 2006; Behrens et al., 2007; Jocham et al., 2009b). Although our task required integration of outcome history, the VLSM analysis did not identify any voxels around dACC. We had good coverage of dACC in our sample, and the power analysis confirms that we could detect the effect of lesions in this area. Furthermore, lesions to this area in the same sample of patients did disrupt performance on other tasks (Tsuchida and Fellows, 2009). One possibility is that intact dACC is more critical when outcome history is learned through action–outcome association rather than stimulus–outcome association (Ostlund and Balentine, 2007; Rushworth et al., 2007b; Rudebeck et al., 2008).

In previous lesion work, a subgroup of patients with damage affecting DLPFC was found to perform as poorly as OFC patients on a probabilistic reversal learning task (Hornak et al., 2004). Based on the posttest debriefing, the authors concluded that this impairment was attributable to a failure to attend to the essential feedback information rather than impaired reversal learning per se. Single-unit work in monkeys also suggests that neurons in DLPFC carry information about appropriate strategies in feedback-driven learning (Genovesio et al., 2005). In the present study, LF patients as a group showed an intermediate pattern of performance, not significantly different from either OFC or healthy control groups, reflecting variable performance within the group. Note also that we had limited power to detect effects related to left lateral frontal damage in this sample. The VLSM analysis suggests that impairment in the LF group may relate to damage encroaching on the most lateral aspect of OFC rather than injury to the frontal convexity proper, but additional work is needed to definitively address this issue.

In summary, our findings demonstrate the critical contribution of OFC, but not other areas within PFC, to both initial learning and reversal learning in a probabilistic environment. The pattern of impairment observed in patients with OFC damage suggests that OFC does not simply function to signal outcome information or to suppress an existing stimulus–outcome association in a changing environment but instead is critical for strategically applying outcome information in a context-dependent manner in uncertain environments.

References

- Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.
- Berlin HA, Rolls ET, Kischka U (2004) Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain* 127:1108–1126.
- Brunner E, Munzel U (2000) The nonparametric Behrens–Fisher problem: asymptotic theory and a small-sample approximation. *Biom J* 42:17–25.

- Budhani S, Marsh AA, Pine DS, Blair RJ (2007) Neural correlates of response reversal: considering acquisition. *Neuroimage* 34:1754–1765.
- Chamberlain SR, Müller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ (2006) Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science* 311:861–863.
- Chudasama Y, Passeti F, Rhodes SE, Lopian D, Desai A, Robbins TW (2003) Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: differential effects on selectivity, impulsivity and compulsivity. *Behav Brain Res* 146:105–119.
- Cools R, Clark L, Owen AM, Robbins TW (2002) Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 22:4563–4567.
- Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380:69–72.
- Fellows LK, Farah MJ (2003) Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain* 126:1830–1837.
- Folstein MF, Folstein SE, McHugh PR (1975) “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189–198.
- Genovesio A, Brasted PJ, Mitz AR, Wise SP (2005) Prefrontal cortex activity related to abstract response strategies. *Neuron* 47:307–320.
- Gläscher J, Hampton AN, O’Doherty JP (2009) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* 19:483–495.
- Hampton AN, Adolphs R, Tyszka MJ, O’Doherty JP (2007) Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron* 55:545–555.
- Hornak J, O’Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE (2004) Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J Cogn Neurosci* 16:463–478.
- Iversen SD, Mishkin M (1970) Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 11:376–386.
- Jocham G, Klein TA, Neumann J, von Cramon DY, Reuter M, Ullsperger M (2009a) Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. *J Neurosci* 29:3695–3704.
- Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M (2009b) Adaptive coding of action values in the human rostral cingulate zone. *J Neurosci* 29:7489–7496.
- Jones B, Mishkin M (1972) Limbic lesions and the problem of stimulus–reinforcement associations. *Exp Neurol* 36:362–377.
- Kahnt T, Park SQ, Cohen MX, Beck A, Heinz A, Wrase J (2009) Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *J Cogn Neurosci* 21:1332–1345.
- Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9:940–947.
- Kimberg DY, Coslett HB, Schwartz MF (2007) Power in voxel-based lesion-symptom mapping. *J Cogn Neurosci* 19:1067–1080.
- Murphy FC, Michael A, Robbins TW, Sahakian BJ (2003) Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. *Psychol Med* 33:455–467.
- Murray EA, Izquierdo A (2007) Orbitofrontal cortex and amygdala contributions to affect and action in primates. *Ann NY Acad Sci* 1121:273–296.
- Murray EA, O’Doherty JP, Schoenbaum G (2007) What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J Neurosci* 27:8166–8169.
- Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53:695–699.
- O’Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95–102.
- O’Doherty J, Critchley H, Deichmann R, Dolan RJ (2003) Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J Neurosci* 23:7931–7939.
- Ostlund SB, Balleine BW (2007) Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J Neurosci* 27:4819–4825.
- Remijne PL, Nielen MM, Uylings HB, Veltman DJ (2005) Neural correlates of a reversal learning task with an affectively neutral baseline: an event-related fMRI study. *Neuroimage* 26:609–618.
- Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS (2004) Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn* 56:129–140.
- Rolls ET (2004) The functions of the orbitofrontal cortex. *Brain Cogn* 55:11–29.
- Rolls ET, Hornak J, Wade D, McGrath J (1994) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J Neurol Neurosurg Psychiatry* 57:1518–1524.
- Rorden C, Brett M (2000) Stereotaxic display of brain lesions. *Behav Neurol* 12:191–200.
- Rorden C, Karnath HO, Bonilha L (2007) Improving lesion-symptom mapping. *J Cogn Neurosci* 19:1081–1088.
- Rorden C, Fridriksson J, Karnath HO (2009) An evaluation of traditional and novel tools for lesion behavior mapping. *Neuroimage* 44:1355–1362.
- Rudebeck PH, Murray EA (2008) Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *J Neurosci* 28:8338–8343.
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF (2008) Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 28:13775–13785.
- Rudrauf D, Mehta S, Bruss J, Tranel D, Damasio H, Grabowski TJ (2008) Thresholding lesion overlap difference maps: application to category-related naming and recognition deficits. *Neuroimage* 41:970–984.
- Rushworth MF, Buckley MJ, Behrens TE, Walton ME, Bannerman DM (2007a) Functional organization of the medial frontal cortex. *Curr Opin Neurobiol* 17:220–227.
- Rushworth MF, Behrens TE, Rudebeck PH, Walton ME (2007b) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn Sci* 11:168–176.
- Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M (2003) Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learn Mem* 10:129–140.
- Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK (2009) A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat Rev Neurosci* 10:885–892.
- Swainson R, Rogers RD, Sahakian BJ, Summers BA, Polkey CE, Robbins TW (2000) Probabilistic learning and reversal deficits in patients with Parkinson’s disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia* 38:596–612.
- Tsuchida A, Fellows LK (2009) Lesion evidence that two distinct regions within prefrontal cortex are critical for n-back performance in humans. *J Cogn Neurosci* 21:2263–2275.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, Rushworth MF (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* 65:927–939.
- Wheeler EZ, Fellows LK (2008) The human ventromedial frontal lobe is critical for learning from negative feedback. *Brain* 131:1323–1331.
- Williams ZM, Bush G, Rauch SL, Cosgrove GR, Eskandar EN (2004) Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat Neurosci* 7:1370–1375.