

Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm

Lesley K. Fellows and Martha J. Farah

Center for Cognitive Neuroscience, University of Pennsylvania, Philadelphia, USA

*Correspondence to: Lesley K. Fellows, Center for Cognitive Neuroscience, University of Pennsylvania, 3815 Walnut St, Philadelphia, PA 19104–6196, USA
E-mail: lesleyf@psych.upenn.edu*

Summary

How do the frontal lobes support behavioural flexibility? One key element is the ability to adjust responses when the reinforcement value of stimuli change. In monkeys, this ability—a form of affective shifting known as reversal learning—depends on orbitofrontal cortex. The present study examines the anatomical bases of reversal learning in humans. Subjects with

lesions of the ventromedial prefrontal cortex were compared with a group with dorsolateral frontal lobe damage, as well as with normal controls on a simple reversal learning task. Neither form of frontal damage affected initial stimulus–reinforcement learning; ventromedial frontal damage selectively impaired reversal learning.

Keywords: frontal lobes; ventromedial prefrontal cortex; lesion; decision making; reward

Abbreviations: BDI = Beck Depression Inventory; CTL = age- and education-matched control subjects; DLF = dorsolateral frontal; fMRI = functional MRI; IADL = Instrumental Activities of Daily Living; MNI = Montreal Neurological Institute; VLF = ventrolateral prefrontal; VMF = ventromedial frontal

Introduction

An essential component of intelligent behaviour is the ability to learn the reward value of stimuli, which can change according to an animal's circumstances and state. A previously rewarded stimulus may cease to be rewarding or even reverse its value and become punishing as a function of external changes in reinforcement contingencies or internal changes in appetite and satiety. Single cell recording and lesion studies have demonstrated that the orbitofrontal cortex is the substrate of flexible encoding of stimulus reward value in macaques. Orbitofrontal neurons encode the context-specific reward value of stimuli (Tremblay and Schultz, 1999; Rolls, 2000). Orbitofrontal lesions in monkeys impair 'affective shifting'—the ability to adapt associative learning when an initially rewarded stimulus is no longer rewarded (extinction) or when the reward and punishment value of two stimuli switch (reversal learning). In contrast, lesions of dorsolateral prefrontal cortex do not affect either of these forms of learning (Jones and Mishkin, 1972; Dias *et al.*, 1996).

Surprisingly little is known about the corresponding neural system in humans. Since the time of Phineas Gage, damage to ventral prefrontal cortex has been observed to result in

personality changes, impaired impulse control and alterations in emotional and motivational state (Mesulam, 2002). It has been suggested that some of these changes are related to fundamental alterations in flexible stimulus–reward learning. Two recent lines of research have begun to investigate aberrant reinforcement processing in patients with ventral prefrontal damage, albeit from different perspectives. The first has focused on characterizing the impulsive decision-making that may follow ventromedial frontal (VMF) damage. It has been reported that patients with VMF lesions performing a complex gambling task are unable to change response strategies in the face of accumulating evidence that an initially preferred choice is detrimental in the long run (Bechara *et al.*, 1997, 2000). Using simpler tasks that directly tested reversal learning and extinction, but in less well-characterized patients, Rolls and colleagues have argued more generally that the socially inappropriate behaviour of some frontally-damaged patients may reflect a basic inability to modify on-going behaviour in response to negative feedback (Rolls *et al.*, 1994). Remarkably, however, functional imaging studies have yet to demonstrate a role for orbitofrontal cortex in the flexible stimulus–reinforcement

Table 1 Subject characteristics

Group	n	Age (years)	Sex (F/M)	Education (years)	NART IQ	BDI score	IADL score
VMF	8	57 (12)	4/4	12.9 (2.0)	115 (10)	8.9 (8.1)	17.8 (3.4)
DLF	12	62 (10)	7/5	15.6 (2.7)	118 (11)	9.1 (3.3)	20.4 (1.4)
CTL	12	56 (16)	7/5	14.8 (2.7)	118 (12)	3.8 (2.6)	n/a

See text for details. Mean given with SD in parentheses. NART = National Adult Reading Test.

learning captured by reversal learning tasks; nor has reversal learning been assessed directly in humans with well-defined lesions of the ventromedial frontal (VMF) cortex.

An early study of this kind of stimulus–reward learning in brain-damaged humans—and the most informative to date—was carried out by Rolls and colleagues (Rolls *et al.*, 1994). They examined the relationship between changes in day-to-day behaviour and reversal learning abilities in a heterogeneous group of patients. Subjects with primarily ventral frontal damage were impaired at both extinction and reversal learning, and had more severe disturbances of behaviour, when compared with those with damage outside this area. However, the anatomical specificity of this study was limited: many of the subjects did not have exclusive ventral frontal damage and the majority of the control group had non-frontal damage. This work therefore establishes the dissociability of reversal learning as a unique, frontally-mediated form of learning in humans. It does not, however, tell us what systems of prefrontal cortex participate in this process nor confirm the dorsal–ventral dissociation evident in animal work.

Two functional imaging studies have examined the neural basis of reversal learning in humans. An event-related functional MRI (fMRI) study of probabilistic reversal learning found activation in ventrolateral frontal (VLF) cortex when reversal error trials that preceded a correct response were contrasted with correct trials. Signal loss from the orbitofrontal cortex meant that its contribution to reversal learning could not be examined (Cools *et al.*, 2002). A PET study with better ability to image orbitofrontal cortex observed no activation in this area when reversal learning was contrasted with other phases of a visual discrimination associative learning task (Rogers *et al.*, 2000). More generally, however, several studies have implicated orbitofrontal cortex in reward processing, particularly in the context of uncertain or changing contingencies (Rogers *et al.*, 1999; Elliott *et al.*, 2000; O’Doherty, *et al.*, 2000, 2001).

Technical limitations often make this region of the brain challenging to evaluate with standard fMRI techniques. Further, unlike lesion studies, imaging studies do not allow inferences to be drawn about whether a brain area has a necessary role in a given cognitive process. Therefore, we assess the neural substrates of flexible reinforcement processing in a study of human subjects with focal lesions of prefrontal cortex. Specifically, we test the hypothesis that injury restricted to the VMF lobes, but not to the dorsolateral

frontal (DLF) lobes, impairs reversal learning in humans. We also examined the relationship between reversal learning performance and degree of functional impairment in subjects with frontal lobe damage in an effort to understand the influence of impaired affective shifting on day-to-day behaviour.

Methods

Subjects

Subjects with fixed lesions involving the frontal lobes were identified through the research databases of the Hospital of the University of Pennsylvania and the Moss Rehab Research Institute. Any patient with a lesion confined to the frontal lobes, regardless of laterality, size or aetiology, was eligible. Patients were further classified as having a lesion that primarily involved either the VMF or DLF lobes (see below).

Exclusion criteria included past or intercurrent medical or neurological disease likely to impair cognition. Subjects were tested a minimum of 6 months after the brain injury. Age- and education-matched control subjects (CTL), free of neurological or psychiatric disease and not taking psychoactive medications, were recruited by advertisement in the local community. All participants provided written, informed consent in accordance with the Declaration of Helsinki; the study protocol was approved by the Institutional Review Boards of both participating centres.

Background information is provided in Table 1. There was no significant difference in age, level of education or pre-morbid IQ estimated by the National Adult Reading Test across the three groups of subjects. ANOVA (analysis of variance) demonstrated a significant effect of group on Beck Depression Inventory (BDI) scores [$F(2,28) = 4.4, P = 0.02$]. *Post hoc* Newman–Keuls tests showed that both patient groups had significantly higher scores than controls, but there was no significant difference between the two patient groups.

Four of eight subjects in the VMF group and five of 12 subjects with DLF damage were taking at least one psychoactive medication. These included specific serotonin uptake inhibitors, phenytoin, centrally-acting acetylcholinesterase inhibitors and methylphenidate.

Level of day-to-day function was assessed in two ways. Subjects with frontal damage (or, in three cases, their spouse or parent) completed a standard Instrumental Activities of Daily Living (IADL) scale (Gallo *et al.*, 2000) to assess their day-to-day level of function. The scale was not completed for

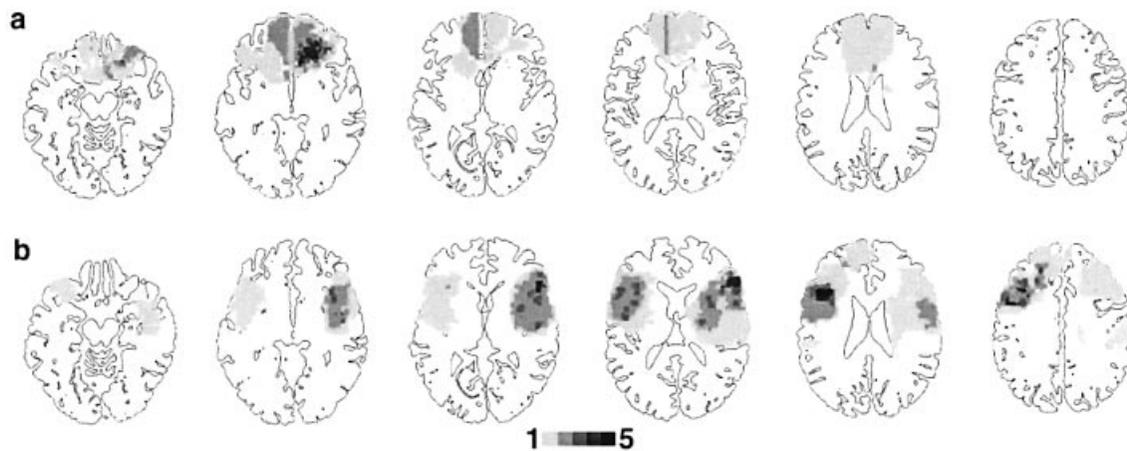


Fig. 1 Location and overlap of brain lesions. (a) Lesions of the eight subjects with ventromedial frontal damage. (b) Lesions of the 12 subjects with DLF damage, projected on the same six axial slices of the standard MNI brain and oriented according to radiological convention. Areas damaged in only one subject are shown in light grey; darker shades denote the degree to which lesions involve the same structures in multiple subjects. This figure is available in colour as supplementary material at Brain Online.

two subjects in the DLF group and one in the VMF group due to administrative problems. This scale evaluates seven domains, such as the ability to use the telephone and to manage finances. Total scores can range from 7 (dependent in all domains) to 21 (fully independent). Although this scale asks about very concrete day-to-day activities, patients with frontal damage may lack insight into their deficits. In order to corroborate the results of this functional self-assessment, a psychologist who knew the subjects well in her capacity as database manager (telephone contacts, arranging research visits, performing screening evaluations and, in several cases, visiting the subjects in their homes), but who was blind to the results of this study, independently classified subjects into one of three functional groups: (i) dependent on at least some IADLs; (ii) impaired, but not dependent; and (iii) independent, functioning at or very near pre-morbid level. The agreement in rank order between the two methods of measuring function was very good (Spearman $\rho = 0.84$, $P = 0.0003$). Given its greater range, the IADL score was adopted as the measure of function in further analyses. The VMF group was more impaired than the DLF group, as measured by the IADL score (Mann–Whitney U test, $P = 0.02$).

Lesion analysis

Lesions were traced from the available CT or MRI onto the standard Montreal Neurological Institute (MNI) brain using MRICro software (Rorden and Brett, 2000). The subjects with frontal lobe injury were divided into two groups based on whether each lesion primarily involved the VMF or DLF lobes. Figure 1 shows the location and degree of overlap of the lesions in the two frontal groups. Lesions were due to rupture of anterior communicating aneurysms in seven out of eight VMF subjects. One subject had suffered an ischaemic

stroke in the territories of both anterior cerebral arteries. DLF lesions were due to ischaemic stroke in eight cases, haemorrhagic stroke in three and tumour resection with local radiotherapy in one.

The lesions of the VMF group overlap maximally in the posteromedial orbitofrontal cortex and subjacent white matter. A region of interest that encompassed this area was outlined on the normalized brain on the basis of anatomical landmarks to permit more fine-grained correlational analysis. This region included the posterior half of the gyrus rectus, medial and middle orbital gyrus bilaterally, corresponding to the posteromedial division of the orbitofrontal cortex according to the nomenclature of Hof (Hof *et al.*, 1995).

A second analysis was performed to examine further the relationship between focal frontal damage and reversal learning performance, avoiding the constraint of a pre-defined region of interest. A lesion overlap image was constructed in which lesions were weighted according to the reversal learning performance of each subject, allowing direct visualization of the structure–function relationship.

Half the patients with dorsolateral involvement had lesions that extended to varying degree into the VLF cortex. Given the fMRI results implicating this region in reversal learning (Cools *et al.*, 2002), the influence of such damage on both initial and reversal learning was also examined. A VLF region of interest was defined on the basis of these fMRI results using the following stereotaxic coordinates: $x = \pm 26$ to ± 50 ; $y = +16$ to $+24$; and $z = -9$ to 8 . The six DLF subjects with associated ventrolateral damage had a mean total lesion volume of 23.5 cc (SD 19), with a mean VLF involvement of 1.0 cc (SD 1.6). Five of the VMF subjects had very minor involvement of the VLF cortex [mean total lesion volume in this sub-group 14 cc (SD 10), mean ventrolateral involvement 0.1 cc (SD 0.1)].

Task

Stimulus–reinforcement association learning and reversal learning were assessed by means of a simple computerized card game with play money stakes. Subjects were dealt two cards at a time from packs of different colours. One pack consistently concealed a \$50 win, the other a \$50 loss. Feedback was provided after each trial. After the learning criterion of eight consecutive cards chosen from the winning pack was met, the contingencies were switched. This constituted the reversal phase of the task. If the criterion was again met, the contingencies were switched again for a total of 50 trials, allowing up to five reversals.

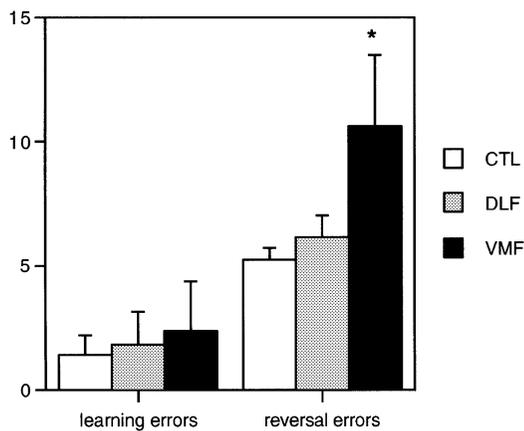


Fig. 2 Initial stimulus–reinforcer association learning and reversal learning performance in subjects with fixed damage of VMF lobes, DLF lobes and CTL. Initial learning performance is expressed as the mean number of errors made before the learning criterion was met and reversal learning as the mean number of errors in the reversal phase of the experiment. Error bars show the 95% confidence intervals.

Statistical analysis

The effect of group membership on errors was examined by repeated measures ANOVA, with learning phase (initial, reversal) the within subjects measure. *Post hoc* pairwise comparisons were made with the Newman–Keuls test. Correlations between continuous variables were assessed with Pearson’s *r*. Spearman rank-order correlations were determined for ordinal data. The significance level was set at $P < 0.05$.

Results

We tested stimulus–reinforcement associative learning and reversal learning in human subjects with damage either to the VMF lobes or DLF lobes, in addition to age- and education-matched control subjects. Errors were submitted to repeated measures ANOVA with group as the between subjects factor, and learning phase (initial, reversal) as the within subjects factor. This revealed significant main effects of group [$F(2,29) = 14.0, P < 0.0001$] and of learning phase [$F(1,29) = 121.1, P < 0.0001$], as well as a significant group \times learning phase interaction [$F(2,29) = 7.3, P = 0.003$]. As shown in Fig. 2, there was no significant difference across groups in the number of errors made while learning the initial discrimination [$F(2,29) = 0.61, P = 0.55$]. There was, however, a significant effect of group membership on the number of errors made when reversal learning was required [$F(2,29) = 20.2, P < 0.0001$]. Subjects with VMF damage made more than twice as many errors as normal controls in the reversal phase of the task. In contrast, the performance of subjects with DLF damage was indistinguishable from controls and significantly different from VMF performance (Newman–Keuls test, $P < 0.05$).

Of the measured demographic characteristics, only education was correlated with reversal learning performance across all subjects ($r = -0.42, P = 0.02$). The ANOVA examining the relationship between errors in the two learning conditions

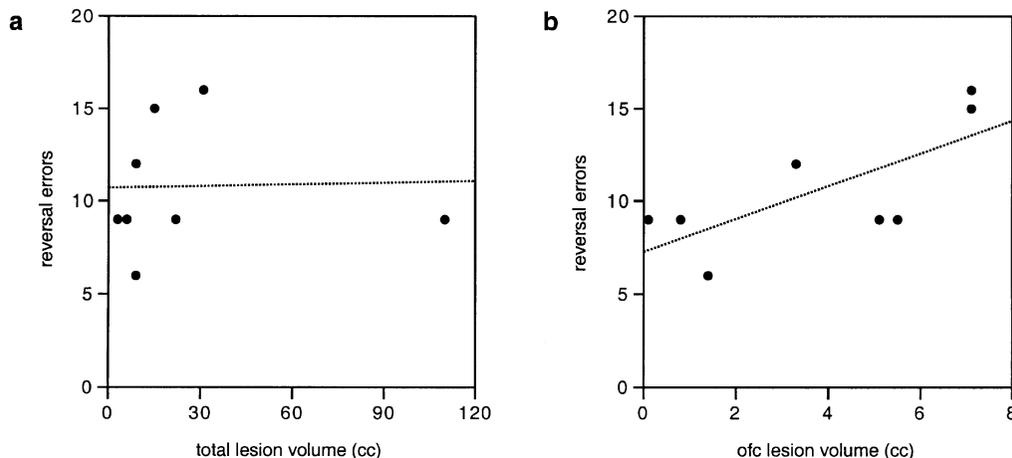


Fig. 3 Correlation between reversal learning performance and volume of damaged tissue in the VMF group. (a) Correlation with total tissue damage. (b) Correlation with volume of tissue damaged within the posteromedial orbitofrontal cortex (ofc). Further details are provided in the text.

(initial, reversal) and group membership (CTL, DLF, VMF) was therefore repeated with level of education, dichotomized as high school or less (≤ 12 years) or any college (>12 years) included as a factor. The results were not substantially altered: effect of group [$F(2,26) = 11.5, P = 0.0003$]; effect of learning condition [$F(1, 26) = 83.3, P < 0.0001$]; effect of group \times learning condition [$F(2,26) = 6.2, P = 0.006$]; and effect of education [$F(1,26) = 1.6, P = 0.22$].

There was no significant correlation between the total volume of damaged tissue and reversal learning performance in the subjects with VMF damage ($r = 0.06, P = 0.81$) nor between lesion volume and reversal learning performance across all subjects with frontal damage ($r = 0.08, P = 0.75$). In contrast, the extent of damage to the posteromedial orbitofrontal cortex and its overlying white matter, a region that included the site of maximum lesion overlap, showed a

significant positive correlation with degree of reversal learning impairment ($r = 0.72, P = 0.04$) (Fig. 3).

In order to evaluate further the relationship between structural damage and reversal learning, a second lesion overlap image was generated for the seven of the eight VMF subjects with abnormal reversal learning performance. These seven subjects also had the worst performance of the frontal group as a whole (≥ 9 errors in the reversal phase). They were divided into three groups on the basis of performance: 9 errors ($n = 4$); 12 errors ($n = 1$); or 15–16 errors ($n = 2$). Individual lesions were then given a weight of 1, 2 or 3 accordingly. The maximal overlap involved 10 of the 12 available layers, and occurred in the white matter in the posteromedial orbitofrontal cortex within a 16 mm depth from the orbitofrontal cortical surface (Fig. 4). Although the area of maximal overlap is on the left, there were too few subjects with unilateral right VMF damage to allow definite conclusions about laterality to be drawn.

The DLF subjects as a group had lesions scattered throughout the DLF lobes, with somewhat less overlap than the VMF group as a consequence. This raises the possibility that the normal reversal learning of the DLF group as a whole is masking an impaired subgroup with damage to a putative key DLF region. As mentioned, when reversal learning performance is examined on a case-by-case basis, the seven worst performers were VMF subjects. Only one DLF subject performed worse than any control, by a margin of a single error. This absence of outliers from the DLF group adds further support to the contention that damage to DLF cortex does not impair reversal learning.

Although this study was not originally designed to evaluate the role of VLF cortex in reversal learning, several subjects in the DLF group had lesions that extended into this region. We therefore performed a *post hoc* region of interest-based analysis to examine whether damage to the area of VLF cortex implicated in reversal learning studied with fMRI (Cools *et al.*, 2002) influenced either initial or reversal learning performance. Table 2 shows the initial and reversal learning performance split by VLF involvement. Although data are presented for the VMF group split in this way, it should be noted that the degree of VLF involvement in the VMF group was minor (see Methods). The subgroups are small, but it does not appear that unilateral damage to VLF influences reversal learning. The only values for which the 95% confidence intervals do not overlap are the error rates for the DLF groups with and without accompanying VLF damage; the presence of VLF damage in the DLF group is associated with better initial learning performance. Three of the six subjects with damage that involved ventrolateral, but not ventromedial, prefrontal cortex had lesions of the right hemisphere; there was no indication of an effect of the laterality of VLF damage on either initial or reversal learning (data not shown).

Depression is a frequent sequela of damage to the frontal lobes (Robinson *et al.*, 1984; Mayberg, 2001), and it has been argued that impaired decision making following frontal

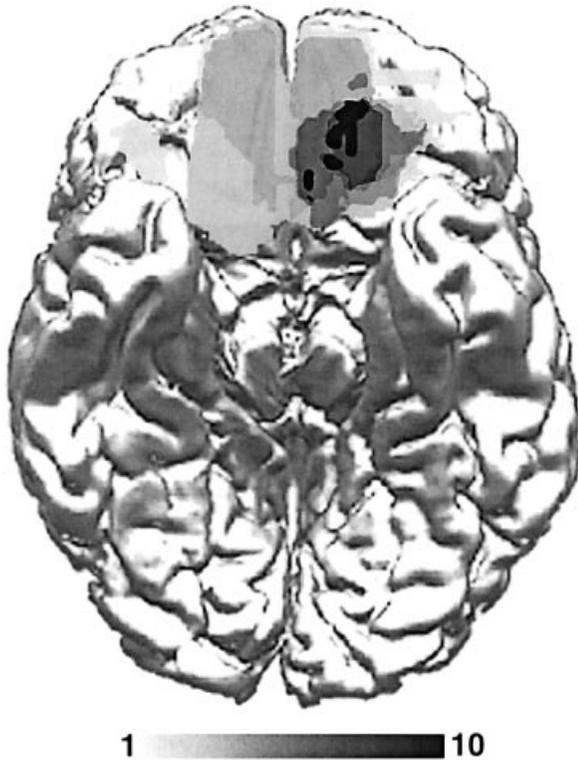


Fig. 4 Neuroanatomical–functional relationship for reversal learning performance. A lesion overlap image was developed for the seven frontal subjects with the worst reversal learning performance (errors ≥ 9), all of whom had VMF damage. The lesion of each subject was weighted by the reversal learning performance of that subject in order to highlight the areas of damage in proportion to their impact on performance (see text). The area of maximal overlap occurred within 16 mm of the orbitofrontal cortical surface. This image is a ‘glass brain’ projection of the overlap within this 16 mm volume shown on the ventral surface of a rendering of the MNI brain, oriented according to radiological convention. The lightest grey represents one functional overlap level, black the overlap of 10 (of a possible 12) functional levels. The area of maximal overlap involves the left posteromedial orbitofrontal cortex. This figure is available in colour as supplementary material at Brain Online.

Table 2 Mean performance on initial stimulus-reinforcement associative learning, and on reversal learning in DLF and VMF groups, split by the presence or absence of additional damage to the VLF

Group	Learning errors mean (95% CI)	Reversal errors mean (95% CI)
DLF not VLF ($n = 6$)	3.3 (1.2–5.4)	6.0 (4.5–7.5)
DLF and VLF ($n = 6$)	0.3 (0–0.87)	6.0 (4.4–7.6)
VMF not VLF ($n = 3$)	4.0 (1.1–6.9)	9.0 (9.0–9.0)
VMF and VLF ($n = 5$)	1.4 (0–6.1)	11.6 (6.4–16.8)

CI = confidence interval.

damage may be due, in part, to co-morbid depression (Murphy *et al.*, 2001; Manes *et al.*, 2002). Although only one subject's BDI scores fell in the range associated with clinical depression (>16), we examined whether mild depressive symptoms were correlated with worse reversal learning performance in the group as a whole. We were unable to detect any such pattern in the present data: there was no correlation between score on the BDI and reversal learning performance in the combined frontal group ($r = -0.06$, $P = 0.81$). Because BDI scores may be increased following stroke for reasons other than depression *per se* (e.g. weight gain due to inactivity, concern about appearance due to hemiparesis), we also divided the subjects into those with any history of depression and those without. Only one VMF subject had a history of depression, which predated his brain injury by many years. In contrast, five out of 12 DLF subjects also carried a diagnosis of depression, in two cases predating their strokes. There was no difference in reversal learning performance in those subjects with a history of depression and those without [$F(1,17) = 0.2$, $P = 0.66$].

Do reversal learning impairments underpin changes in behaviour that are detectable in every day life? Level of function in all subjects with frontal damage was quantified with a standard IADL scale. In the 17 subjects with frontal damage for whom complete data were available, there was a correlation between reversal learning performance and day-to-day functional abilities (Spearman $\rho = -0.57$, $P = 0.01$). In contrast, there was no correlation between total lesion volume and IADL score ($\rho = -0.15$, $P = 0.55$)

Discussion

This study is the first to demonstrate that intact ventromedial prefrontal cortex, but not dorsolateral prefrontal cortex, is necessary for normal reversal learning in humans. This is not a general associative learning deficit: neither group of frontal patients was significantly worse than controls at learning the initial association in this simple task.

Although several of the subjects in the VMF group had damage extending into the medial frontal lobes in addition to orbitofrontal cortex involvement, correlational analysis indicates that the key area of damage leading to impaired

reversal learning is the medial orbitofrontal cortex. More extensive damage to adjacent ventral and/or medial frontal structures, as quantified by the total lesion volume, did not correlate with worse performance. At least one selective lesion study in macaques has also shown that sustained impairments in reversal learning are specific to medial orbitofrontal cortex lesions (Iversen and Mishkin, 1970), although this finding has not been universal (Butter, 1969).

Similarly, additional damage extending into VLF cortex did not adversely affect either initial learning or reversal learning performance in either the VMF or DLF group. This finding contrasts with the fMRI study of Cools *et al.* (2002), in which right VLF cortex activation was found to be associated with reversal error trials. This VLF activity may be related to the additional demands of the more complex probabilistic reversal learning task used in that study. Alternatively, the putative role of this VLF region in reversal learning may be bilaterally represented, so that subjects with unilateral lesions remain unimpaired.

Orbitofrontal cortex dysfunction has been hypothesized to play a role in a variety of neurological and psychiatric conditions, including fronto-temporal dementia (Rahman *et al.*, 1999), addiction (Volkow and Fowler, 1999), depression (Drevets *et al.*, 2001) and obsessive-compulsive disorder (Nielen *et al.*, 2002), to name a few. Behavioural methods to study the role of this frontal area in humans are limited; although VMF lobe damage impairs performance on a widely used gambling task (Bechara *et al.*, 1997), recent work has called the anatomical specificity of this effect into question (Manes *et al.*, 2002). Reversal learning tasks may provide useful tools for selectively probing orbitofrontal cortex function in such patient populations.

Lesion work in animals has demonstrated double dissociations between reversal learning (also termed affective shifting) and the ability to learn associations between reinforcement and new stimulus features for which the suppression of previous stimulus-reinforcement associations is not required (attentional shifting). Ventral prefrontal lesions impair affective shifts, but leave attentional shifts intact, while dorsal prefrontal lesions do the opposite (Dias *et al.*, 1996). The concomitant experiment has yet to be reported in humans. Establishing the existence of such double dissociations would add considerably to the limited human literature in which lesion methods have been employed to study the separate processes mediated by dorsal and ventral prefrontal cortex (Bechara *et al.*, 1998; Manes *et al.*, 2002). The present work demonstrates a single dissociation for affective shifting, supporting the contention that this framework for understanding the anatomical basis of stimulus-reinforcement learning in animals is also of relevance in humans.

A previous study of reversal learning in humans with frontal lobe damage reported a correlation between reversal learning performance across all subjects with frontal damage and a global rating of the severity of behavioural change (Rolls *et al.*, 1994). We also found that level of function—in

this case measured with a simple IADL scale—was negatively correlated with reversal learning impairment. Both these pieces of evidence raise the possibility that impaired affective shifting manifests itself in functionally important ways in everyday human behaviour, although experiments that identify the affected behaviours in more detail, and that move beyond correlational evidence, are required to explore this issue further.

The existence of neural circuitry specialized for rapidly unlearning or suppressing the influence of an established stimulus-reinforcement association is counter-intuitive. Our findings, adding to an extensive animal literature, argue that such a mechanism exists in humans and is, at least in part, anatomically distinct from the neural substrates that mediate the initial learning process.

Acknowledgements

We wish to thank Dr Marianna Stark for her help with subject recruitment and functional assessment. We also wish to thank the anonymous reviewers for their helpful commentary. This research is supported by NIH grants R01-AG14082, R01-DA14129, and NSF grant no. 0226060. L.K.F is supported by a Clinician-Scientist award from the Canadian Institutes of Health Research.

References

- Bechara A, Damasio H, Tranel D, Damasio AR. Deciding advantageously before knowing the advantageous strategy. *Science* 1997; 275: 1293–5.
- Bechara A, Damasio H, Tranel D, Anderson SW. Dissociation of working memory from decision making within the human prefrontal cortex. *J Neurosci* 1998; 18: 428–37.
- Bechara A, Tranel D, Damasio H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 2000; 123: 2189–202.
- Butter CM. Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in *Macaca mulatta*. *Physiol Behav* 1969; 4: 163–71.
- Cools R, Clark L, Owen AM, Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 2002; 22: 4563–7.
- Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 1996; 380: 69–72.
- Drevets WC. Neuroimaging and neuropathological studies of depression: implications for the cognitive-emotional features of mood disorders. [Review]. *Curr Opin Neurobiol* 2001; 11: 240–9.
- Elliott R, Dolan RJ, Frith CD. Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb Cortex* 2000; 10: 308–17.
- Gallo JJ, Fulmer T, Paveza GJ, Reichel W. Handbook of geriatric assessment. 3rd ed. Gaithersburg (MD): Aspen Publications; 2000. p. 110.
- Hof PR, Mufson EJ, Morrison JH. Human orbitofrontal cortex: cytoarchitecture and quantitative immunohistochemical parcellation. *J Comp Neurol* 1995; 359: 48–68.
- Iversen SD, Mishkin M. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 1970; 11: 376–86.
- Jones B, Mishkin M. Limbic lesions and the problem of stimulus-reinforcement associations. *Exp Neurol* 1972; 36: 362–77.
- Manes F, Sahakian B, Clark L, Rogers R, Antoun N, Aitken M, et al. Decision-making processes following damage to the prefrontal cortex. *Brain* 2002; 125: 624–39.
- Mayberg HS. Frontal lobe dysfunction in secondary depression. In: Salloway SP, Malloy PF, Duffy JD, editors. *The frontal lobes and neuropsychiatric illness*. Washington (DC): American Psychiatric Publishing; 2001. p. 167–86.
- Mesulam MM. The human frontal lobes: transcending the default mode through contingent encoding. In: Stuss DT, Knight RT, editors. *Principles of frontal lobe function*. Oxford: Oxford University Press; 2002. p. 8–30.
- Murphy FC, Rubinsztein JS, Michael A, Rogers RD, Robbins TW, Paykel ES, et al. Decision-making cognition in mania and depression. *Psychol Med* 2001; 31: 679–93.
- Nielen MMA, Veltman DJ, de Jong R, Mulder G, den Boer JA. Decision making performance in obsessive compulsive disorder. *J Affect Disord* 2002; 69: 257–60.
- O’Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F, Kobal G, et al. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport* 2000; 11: 893–7.
- O’Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 2001; 4: 95–102.
- Rahman S, Sahakian BJ, Hodges JR, Rogers RD, Robbins TW. Specific cognitive deficits in mild frontal variant frontotemporal dementia. *Brain* 1999; 122: 1469–93.
- Robinson RG, Kubos KL, Starr LB, Rao K, Price TR. Mood disorders in stroke patients. Importance of location of lesion. *Brain* 1984; 107: 81–93.
- Rogers RD, Owen AM, Middleton HC, Williams EJ, Pickard JD, Sahakian BJ, et al. Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *J Neurosci* 1999; 19: 9029–38.
- Rogers RD, Andrews TC, Grasby PM, Brooks DJ, Robbins TW. Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J Cogn Neurosci* 2000; 12: 142–62.
- Rolls ET. The orbitofrontal cortex and reward. [Review]. *Cereb Cortex* 2000; 10: 284–94.
- Rolls ET, Hornak J, Wade D, McGrath J. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J Neurol Neurosurg Psychiatry* 1994; 57: 1518–24.

Rorden C, Brett M. Stereotaxic display of brain lesions. *Behav Neurol* 2000; 12: 191–200.

Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. *Nature* 1999; 398: 704–8.

Volkow ND, Fowler JS. Addiction, a disease of compulsion and

drive: involvement of the orbitofrontal cortex. [Review]. *Cereb Cortex* 2000; 10: 318–25.

Received January 20, 2003. Revised March 6, 2003.

Accepted March 27, 2003.