Reward-related Reversal Learning after Surgical Excisions in Orbito-frontal or Dorsolateral Prefrontal Cortex in Humans

J. Hornak[1], J. O'Doherty[3], J. Bramham[2], E. T. Rolls[1], R. G. Morris[2], P. R. Bullock[4], and C. E. Polkey[4]

Abstract

■ Neurophysiological studies in primates and neuroimaging studies in humans suggest that the orbito-frontal cortex is involved in representing the reward value of stimuli and in the rapid learning and relearning of associations between visual stimuli and rewarding or punishing outcomes. In the present study, we tested patients with circumscribed surgical lesions in different regions of the frontal lobe on a new visual discrimination reversal test, which, in an fMRI study (O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001), produced bilateral orbito-frontal cortex activation in normal subjects. In this task, touching one of two simultaneously presented patterns produced reward or loss of imaginary money delivered on a probabilistic basis to minimize the usefulness of verbal strategies. A number of types of feedback were present on the screen. The main result was that the group of patients with bilateral orbito-frontal cortex lesions were severely impaired at the reversal task, in that they accumulated less money. These patients often failed to switch their choice of stimulus after a large loss and often did switch their choice although they had just received a reward. The investigation showed that bilateral lesions were required for this deficit, since patients with unilateral orbito-frontal cortex (or medial prefrontal cortex) lesions were not impaired in the probabilistic reversal task. The task ruled out a simple motor disinhibition as an explanation of the deficit in the bilateral orbito-frontal cortex patients, in that the patients were required to choose one of two stimuli on each trial. A comparison group of patients with dorsolateral prefrontal cortex lesions was in some cases able to do the task, and in other cases, was impaired. Posttest debriefing showed that all the dorsolateral prefrontal patients who were impaired at the task had failed to pay attention to the crucial feedback provided on the screen after each trial about the amount won or lost on each trial. In contrast, all dorsolateral patients who paid attention to this crucial feedback performed normally on the reversal task. Further, it was confirmed that the bilateral orbito-frontal cortex patients had also paid attention to this crucial feedback, but in contrast had still performed poorly at the task. The results thus show that the orbital prefrontal cortex is required bilaterally for monitoring changes in the reward value of stimuli and using this to guide behavior in the task; whereas the dorsolateral prefrontal cortex, if it produces deficits in the task, does so for reasons related to executive functions, such as the control of attention. Thus, the ability to determine which information is relevant when making a choice of pattern can be disrupted by a dorsolateral lesion on either side, whereas the ability to use this information to guide behavior is not disrupted by a unilateral lesion in either the left or the right orbito-frontal cortex, but is severely impaired by a bilateral lesion in this region. Because both abilities are important in many of the tasks and decisions that arise in the course of daily life, the present results are relevant to understanding the difficulties faced by patients after surgical excisions in different frontal brain regions. ■

INTRODUCTION

There is growing evidence from functional neuroimaging, from the investigation of brain damaged patients, and from neurophysiological and lesion studies in nonhuman primates that the ventral parts of the frontal lobe, which include the orbito-frontal cortex, play a crucial role in representing the reward and punishment value of stimuli and in rapidly learning or reversing associations between visual stimuli and rewards or punishments (Rolls, 1999a, 1999b,2000, 2002). In so far as emotions can be defined as states elicited by rewards and punishments (Rolls, 1990, 1999a, 1999b), these findings may be relevant to understanding the emotional changes and behavioral problems that can follow damage to this region in humans (Hornak, Rolls, & Wade, 1996; Rolls, Hornak, Wade, & McGrath, 1994; Rolls, 1999a).

The importance of the orbito-frontal cortex in processing the reward value of stimuli and in reward-related learning is demonstrated by the finding that single neurons in the primate orbito-frontal cortex respond to the relative reward value of primary reinforcers, such as

[1]University of Oxford, [2]Institute of Psychiatry, De Crespigny Park, [3]Institute of Neurology, Queen's Square, [4]King's Neuroscience Centre

taste (Rolls, Sienkiewicz, & Yaxley, 1989). Other orbitofrontal cortex neurons are involved in one-trial relearning of associations between visual stimuli and a taste reinforcer in the reversal of a visual discrimination task (Rolls, Critchley, Mason, & Wakeman, 1996; Thorpe, Rolls, & Maddison, 1983). A visual discrimination task, and its reversal, allows the learning and reversal of stimulus–reinforcement associations to be investigated (Rolls, 1990, 1999a, 1999b, 2000, 2002). Correspondingly, in human neuroimaging studies, it has been shown that the orbito-frontal cortex is activated by rewarding and punishing stimuli in many different modalities, such as taste (O'Doherty, Rolls, Francis, McGlone, & Bowtell, 2001; Small et al., 1999; Zald, Lee, Fluegel, & Pardo, 1998), odor (Rolls, Kringelbach, et al., 2003; Gottfried, Deichmann, Winston, & Dolan, 2002; O'Doherty et al., 2000; Zald & Pardo, 1997), touch (Rolls, O'Doherty, et al., 2003; Francis et al., 1999), auditory (Frey, Kostopoulos, & Petrides, 2000; Blood, Zatorre, Bermudez, & Evans, 1999), and visual stimuli (O'Doherty et al., 2003; Aharon et al., 2001). Furthermore, neuroimaging has revealed activation of the orbito-frontal cortex to even abstract rewards and punishments, such as monetary gains and losses (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Knutson, Fong, Adams, Varner, & Hommer, 2001; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001b) and verbal feedback (Elliot, Frith, & Dolan, 1997; Elliot, Dolan, & Frith, 2000).

Consistent with these findings, lesion studies in both nonhuman primates and in human patients further indicate the importance of the orbito-frontal cortex in mediating behavioral choice following rewarding or punishing feedback. In visual discrimination tasks, orbito-frontal cortex lesions in nonhuman primates produce impairments at extinguishing or switching responses from a previously rewarded stimulus when the contingencies are altered (Meunier, Bachevalier, & Mishkin, 1997; Iversen & Mishkin, 1970; Butter, 1969). More recently Dias, Robbins, and Roberts (1996) described a dissociation between the effects of lesions in the orbitofrontal cortex, which impaired the ability of monkeys to alter behavior in response to fluctuations in the affective significance of stimuli (in ''intradimensional shift'' conditions) and of lesions in the lateral prefrontal cortex (BA 9), which caused a loss of control of attentional selection (in ''extradimensional shift'' conditions).

Human patients with damage that includes the orbitofrontal cortex have also been shown to have deficits in tasks that involve using rewarding and punishing feedback to guide behavior. Bechara, Damasio, Damasio, and Anderson (1994) reported impairments in ''affective'' decision-making in a gambling task in patients with bilateral ventromedial prefrontal cortex damage (incorporating orbito-frontal, medial prefrontal, and rostral anterior cingulate regions). Damage to these regions impaired the ability to learn to choose advantageously from a set of stimuli that yielded differing overall levels of reward. In an earlier study of visual discrimination reversal learning, Rolls et al. (1994) showed that patients with similar ventromedial lesions were impaired at learning to choose advantageously in a visual discrimination reversal task, in that they repeatedly chose the previously rewarded visual pattern after the contingencies had reversed, a result consistent with a role for the human orbito-frontal cortex in stimulus–reward learning, and specifically in altering behavior in the face of changing reward contingencies. The ventromedial prefrontal lesions in the patients in our earlier study were produced either by closed head injury or by severe strokes, so that the lesions were not restricted just to the ventromedial prefrontal cortex, and in some cases, there was diffuse damage outside the frontal lobe. A major aim of the present study was therefore to determine whether impairments at reversal learning can be produced specifically by lesions of the orbito-frontal cortex. An additional aim was to determine whether unilateral damage to the orbito-frontal cortex can produce this learning impairment or whether it is produced by bilateral damage. We therefore tested patients with lesions that were circumscribed to different regions of the prefrontal cortex because the lesions were produced surgically.

We used a new probabilistic reversal task, developed specifically to minimize the opportunity to use a verbal strategy and to ensure that any impairment on the task could not be attributed to simple motor disinhibition. In this new task, two simple patterns appeared together on a touch-screen on each trial, and selection of each pattern could give and take away varying amounts of imaginary money. Choice of the ''good'' pattern probabilistically gave more than it took overall, whereas the opposite was true of the other ''bad'' pattern. The patient's goal was to determine by trial and error which pattern was more profitable to touch. There was a warning that a reversal would occur, and that this would happen gradually. The task was to keep track of whichever pattern was currently the ''good'' pattern and to keep touching it until the participant thought it had changed and was now the ''bad'' pattern. Since the task itself was difficult enough to avoid ceiling effects in the normal group, it was possible to give explicit instructions explaining that reversals would occur.

In our previous study of reversal (Rolls et al., 1994), only one stimulus (S+ or S−) appeared at a time and the patients had either to make or withhold a response on each trial. The perseverative touching of the old S+ may therefore have had a motor component, an interpretation that would fit with the view of some authors that the orbito-frontal cortex is involved in ''inhibitory control''—a function that could be especially important for inhibiting inappropriate responses during tasks, such as the reversal of visual discrimination learning (Roberts & Wallis, 2000; Dias et al., 1996). Because both stimuli appeared on each trial in the new probabilistic reversal test and because a response was therefore required on

every trial, continued selection of the previously correct stimulus could therefore no longer be open to this motor disinhibition interpretation.

A further rationale for using this new probabilistic reversal test in patients with orbito-frontal cortex damage was that the very same task (save for a number of minor modifications) was used in an fMRI study of visual discrimination reversal learning in healthy normal subjects (O'Doherty, Kringelbach, et al., 2001). It was shown that bilateral activation of the medial orbito-frontal cortex was correlated with the amount on money won on individual trials, and that bilateral activation of the lateral orbito-frontal cortex was correlated with how much money was lost on individual trials. The present study addresses whether these regions of the orbito-frontal cortex where activations were found are necessary for good performance in the visual discrimination reversal task.

In this study, we also included groups of patients with unilateral lesions in other regions of the prefrontal cortex, in particular to the dorsolateral and/or the medial prefrontal cortex. Contrary to the functions ascribed to the orbito-frontal cortex in reward and reward-related learning, the dorsolateral prefrontal cortex has been implicated in different functions. These can be placed under the general umbrella term of ''executive functions'' and include planning (Owen, 1997), working memory (Goldman-Rakic, 1996), higher order response selection (Wise, Murray, & Gerfen, 1996), inhibitory control of attentional selection (Robbins, 1998; Dias et al., 1996), and extradimensional shift learning (Rogers, Andrews, Grasby, Brooks, & Robbins, 2000). One hypothesis that arises from the putatively different roles ascribed to these different portions of the prefrontal cortex is that patients with dorsolateral prefrontal cortex lesions may be less impaired or produce a different pattern of impairments on a task designed to assess reward-related learning and therefore presumed to rely more on orbito-frontal cortex function. Consistent with this, a partial dissociation between the effects of ventromedial and dorsolateral prefrontal cortex was reported by Bechara, Damasio, Tranel, and Anderson (1998) in relation to affective decision making and working memory, respectively. Consequently, a further aim of the present study was to determine whether we could dissociate effects of lesions of the orbito-frontal cortex and dorsolateral prefrontal cortex on the reversal learning task.

RESULTS

All patients succeeded in reaching criterion in the practice session (which did not include a reversal), demonstrating that they had understood the requirements of the task and appreciated the probabilistic nature of the reward schedule. There were no significant differences between the patient groups in the number of trials taken to achieve criterion in the practice session, using a nonparametric analysis of variance (Kruskal–Wallis). (The mean number of trials to attain criterion across all groups was 28.2.) All groups also performed comparably on the reversal test until the contingencies reversed for the first time, gaining approximately £800–1000 over the first 10–13 trials. Thereafter, as shown in Figure 1, all of the patient groups performed like the normal control group with the exception of the bilateral orbito-frontal cortex group and a subset of the unilateral dorsolateral prefrontal cortex patients who, it was found, had failed to attend to the essential feedback (see below). Whereas patients in these two groups lost more and more ''money'' as the test proceeded, the other groups continued to gain. Statistical analyses are presented subsequently.

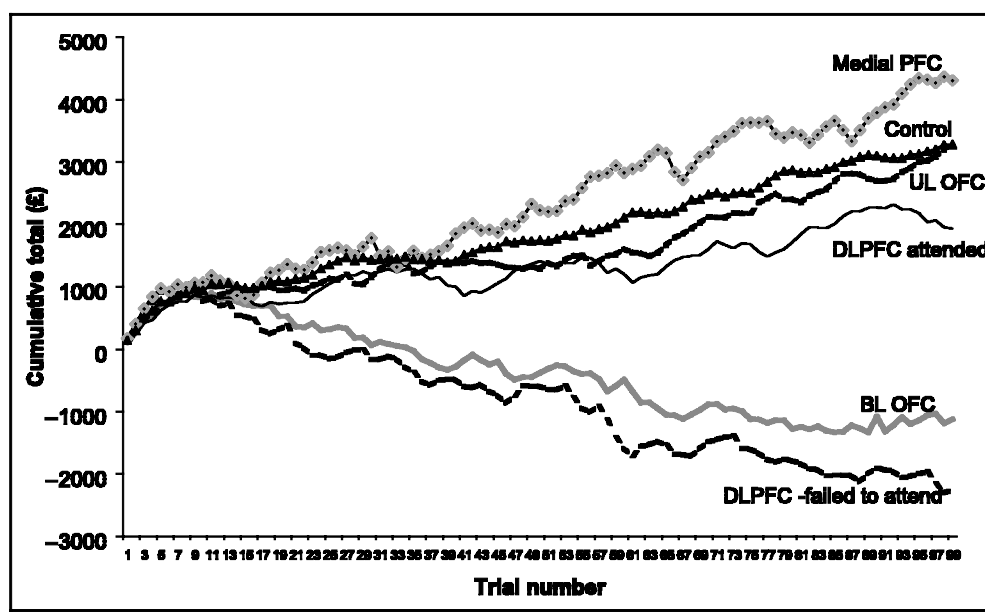

Figure 1. Cumulative total amount of ''money'' won by the normal group and by each of the patient groups by Trial 100. BL OFC = bilateral orbito-frontal cortex; UL OFC = unilateral orbito-frontal cortex; Medial PFC = medial prefrontal cortex; DLPFC = dorsolateral prefrontal cortex. Attended/failed to attend = the patient attended/failed to attend to the crucial feedback during the reversal test, namely, the amount won or lost on each trial.

Posttest Questionnaire

Feedback Found Most Useful

Although all patients were able to use the essential feedback concerning the amount won or lost on each trial during the brief, easy prereversal phase of the test, the results of the posttest questionnaire revealed an important difference between groups in the ability to appreciate which was the essential feedback when the test became more challenging (i.e., during the reversal phase of the test). Thus, 8 of the 17 patients whose lesions included the dorsolateral prefrontal region unilaterally, when asked the open-ended question ''What information on the screen did you find most useful in keeping track of which pattern was currently the good one?'' explained how they had relied on various nonessential sources of feedback and, when asked about the feedback showing the amount won or lost on each trial, denied that they had found this feedback to be the most useful (see Table 1). By contrast, all of the 14 patients whose lesion (bilateral or unilateral) did not encroach on the dorsolateral prefrontal region reported that this was the obvious feedback to use and that they had used it to guide their choice of stimulus. This was also true for all normal control subjects. Statistical comparison between patients with and without a dorsolateral lesion revealed a very significant difference between the numbers in each group who did and who did not attend to the essential feedback (Fisher's exact test, $p = .006$, two-tailed).

Other Feedback Used

Patients with dorsolateral prefrontal damage who failed to attend to the amount won/lost on each trial reported instead that they had used a variety of subsidiary nonessential types of feedback, which, by themselves, would not allow the subject to perform the reversal task successfully (see Appendices 2A and 2B).

Total Scored by Trial 100

Group Results

We tested for a difference between the groups in the total money gained by trial 100. Statistical analysis revealed a highly significant difference (Kruskal–Wallis, $\chi^2(4) = 23.39$, $p < .001$). Paired comparisons between the separate groups and the normal control group, as shown in Table 2, revealed a significant impairment only in the bilateral orbito-frontal cortex group (Mann–Whitney test, $p = .002$, two-tailed) and the dorsolateral (failed to attend) group ($p < .001$, two-tailed). Table 2 also shows the results for the other paired comparisons between the groups. This reveals that the bilateral orbito-frontal cortex group was significantly impaired relative to the unilateral orbito-frontal cortex group ($p = .028$, two-tailed), and that there was no significant difference between the bilateral orbito-frontal cortex group and the unilateral dorsolateral (failed to attend) group. (We note that the two critical comparisons shown in Table 2 between the bilateral orbito-frontal cortex group and normal controls and between the ''dorsolateral prefrontal cortex failed to attend'' group and normal controls are still significant after Bonferroni correction for multiple comparisons.)

Individual Results

The number of standard deviations (*SD*) was calculated by which each patient's total (gained by trial 100) differed from the mean amount won by the Control group by trial 100. Figure 2 and Table 1 show each patient's *SD* in relation to whether they attended or failed to attend to the essential feedback. Most patients with dorsolateral prefrontal lesions who were able to determine which feedback cue they should use and who kept using this feedback throughout the test performed normally (the exceptions being D.B. and R.D.). By contrast, all but one of those who failed to attend to this feedback (namely, R.C.) were severely impaired. The other unilateral groups, those with medial and those with orbital lesions (both of whom attended to the appropriate cues) were both unimpaired. It is important to note that the bilateral orbito-frontal cortex group was impaired although all of the patients had attended to the appropriate feedback. There was no correlation between the total money won by trial 100 and IQ (full-scale IQ, as measured by the National Adult Reading Test, NART) (Spearman's rho = 2.77, $p = .88$). There was also no difference between the IQ of those with dorsolateral lesions who attended to the essential feedback and those who failed to do so (Mann–Whitney $U = 16.5$, $p = .06$).

Appendix 2A shows the responses to questions about the other, nonessential types of feedback. Patients who said that they had used the amount won/lost on each trial were clear about this and generally indicated that they felt that this was the obvious information to use. These patients also noticed/paid attention to other feedback types, but they were less clear about whether they had used this information to guide their choice of pattern. Among those with dorsolateral lesions classified as not having attended to the amount won/lost on each trial, some had done so in an inconsistent and/or intermittent way. Most of those who failed to use the amount won/lost on each trial reported that they had relied upon the numerical total instead. The other three types of feedback were noticed by most patients, but they were not always clear about whether they had been aware of their existence or whether they had been helpful in making a choice.

Lesion Size

There was no correlation between performance on the reversal test (as measured by the total earned

Table 1. Use of Essential Feedback in Relation to Total Earned on the Reversal Test

	Reversal Performance (Number of SDs)	*Attended*	*Failed to Attend*
Without dorsolateral lesion			
Bilateral orbital/medial			
R.F.[a]	−2.4**	+	
V.Z.[a]	−4.9**	+	
V.U.[a]	−5.5**	+	
J.A.[a]	−2.1**	+	
R.R.	0.2	+	
V.O.	−1.3	+	
Unilateral orbital/medial			
R.Q.	1.0	+	
L.J.	−0.9	+	
C.L.	−0.4	+	
T.R.	0.7	+	
S.I. (L)	−0.3	+	
Unilateral medial			
O.F. (L)	1.1	+	
F.G. (L)	−2.7**	+	
E.E.	0.2	+	
With dorsolateral lesion			
U.C.	−0.1	+	
B.R.	−0.1	+	
Q.G.	−0.2	+	
V.F. (L)	0.6	+	
A.R. (L)	−0.3	+	
D.B. (L)	−3.8**	+	
L.S. (L)	−1.5	+	
R.D. (L)	−1.7*	+	
Q.O. (L)	−0.5	+	
R.C.	−1.2		+
F.Z.	−2.0**		+
B.S.	−4.6**		+
G.E	(Gave up)**		+
A.G. (L)	−4.1**		+
L.K. (L)	−6.0**		+

by trial 100) and the size of the lesion, measured as the total area of prefrontal cortex excised (Spearman's rho = −.21, p = .26).

Other Factors: Etiology and Use of Medication

No systematic patterns emerged between type of etiology and performance on the task. Only two patients had suffered from closed head injury: V.O. with a bilateral orbito-frontal cortex lesion and S.I. with a unilateral orbital/medial lesion, and neither of these patients was impaired. It was found that 4 of the 9 patients on medication were impaired (44.4%), compared with 10 of those 23 not on medication (43.5%), indicating that medication did not account for the results.

Location of Lesion in the Bilateral Patients Who Were Impaired

The lesions in the four bilateral patients who were impaired on the reversal test (R.F., V.Z., V.U., and J.A.) all encompassed the regions activated in normal subjects in the parallel fMRI study using the same task (O'Doherty, Kringelbach, et al., 2001), and they spared regions more posterior and lateral to the regions of activation. Conversely, the two unimpaired patients with bilateral lesions (V.O. and R.R.) both had lesions that were anterior and posterior, respectively, to the region activated in the fMRI study.

Other Measures: Staying with a Pattern After a Reward and Switching After a Loss

As a measure of sensitivity to reward, the number of occasions on which patients chose again a pattern from which they had just gained £80 or more was counted. As a measure of sensitivity to punishment, the number of occasions on which patients failed to switch after losses of £250 or more was also counted. Across all patients, both measures were very significantly correlated with the total money won by trial 100 (for reward sensitivity, Spearman's rho = −.59, p = .0018; for punishment sensitivity, Spearman's rho = −.70, p = .0002). Further analysis showed that among those patients with bilateral lesions who were impaired on the reversal, A.S. and V.W. had scores significantly higher than the mean of the control group on both measures, and G.S. only on the

The number of standard deviations by which the patients' total scores differed from the mean from the control group is shown in relation to whether they attended or failed to attend to the amount won or lost on each trial and in relation to the presence of a dorsolateral lesion.

[a]These patients had bilateral lesions that lay within the region activated in the fMRI study in which normal control subjects performed the same reversal task. S.S. and V.O. had lesions outside this region.

*Performed at or below fifth centile.

**Performed at or below first centile.

Table 2. Comparisons of the Amount of Money Earned on the Reversal Test

	BL OFC	*UL OFC*	*DLPFC (Att)*	*DLPFC (F)*
BL OFC				
UL OFC	$p = .028$*			
DLPFC (Att)	$p = .045$*	*ns*		
DLPFC (F)	*ns*	$p = .004$**	$p = .011$*	
Control group	$p = .002$**	*ns*	*ns*	$p < .001$**

Group Comparisons Using Mann–Whitney Nonparametric Tests
Kruskal–Wallis nonparametric one-way ANOVA revealed significant group difference: $\chi^2 = 23.391$, $df = 4$, $p < .001$. BL = bilateral; UL = unilateral; OFC = orbito-frontal cortex; DLPFC = dorsolateral prefrontal cortex; Att = the patients attended to the crucial feedback; F = the patients failed to attend to the crucial feedback.

*Significant at $p < .05$.

**Significant at $p < .01$.

sensitivity to reward measure. B.K. did not have significantly higher scores on either measure, but she made more responses revealing insensitivity to punishment than responses revealing insensitivity to reward.

DISCUSSION

Patients with lesions in different regions of the prefrontal cortex were tested on a probabilistic visual discrimination reversal test in which it was necessary both to determine which feedback was crucial and to use this information appropriately to guide the choice of stimulus to maximize the reward obtained. It was found that bilateral orbital/medial prefrontal lesions produced a severe impairment on this test, whereas even large unilateral lesions, which included the orbital/medial region, had no such effect. It was also found that patients with unilateral medial prefrontal cortex lesions were unimpaired at this task.

The pattern of results was more complicated in the dorsolateral group, all of whom had unilateral lesions, in that some patients were impaired at the task whereas others were not. This was shown to be related to whether, as indicated by the posttest questionnaire, the patients had attended to the feedback necessary to succeed on the task, namely, the amount won or lost on each trial. Nearly half of those with dorsolateral lesions explained how they had attended to other nonessential sources of feedback, attempting to maximize their gains in this way. Those who did attend to the essential feedback performed normally whereas those who did not were as impaired as the group with bilateral orbital lesions.

The pattern of impairment shown by the dorsolateral patients who failed to attend can be contrasted with that shown by the patients with bilateral orbital/medial lesions. Although all of these patients with bilateral orbital/medial prefrontal lesions did attend to the relevant feedback, they were nevertheless unable to use it to adjust their behavior. Instead, when stimulus–reward contingencies were reversed, they frequently chose again the incorrect, previously rewarded stimulus, after large losses or failed to stick with the correct stimulus after gaining money on it. The impairment shown by patients with bilateral orbital/medial lesions is consistent with the findings from our earlier study (Rolls et al., 1994), in which patients with bilateral ventral prefrontal cortex lesions were found to be impaired on a go/no-go

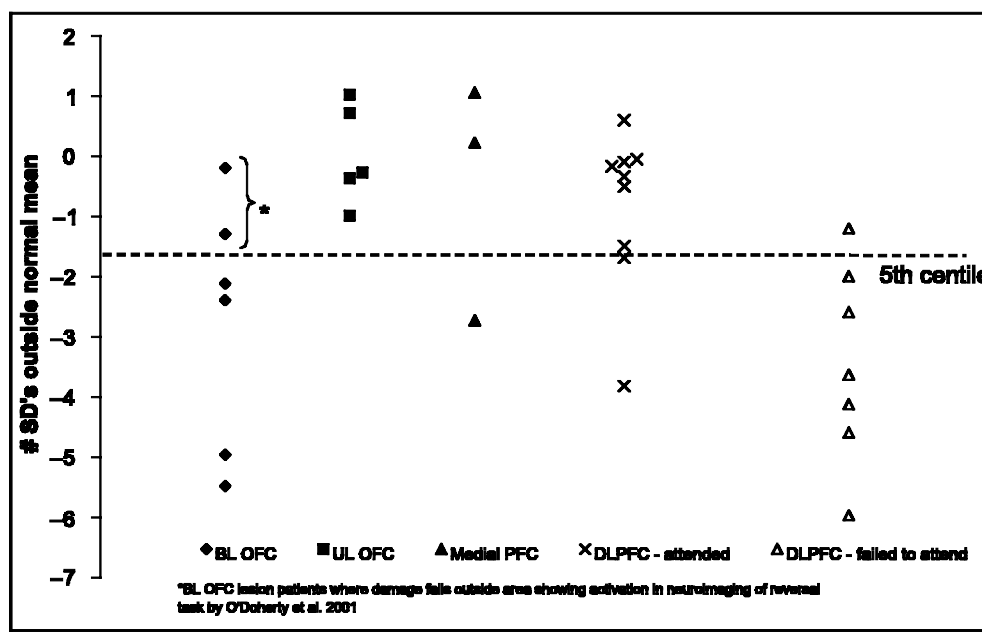

Figure 2. The number of standard deviations by which each patient's total earned on the reversal test by trial 100 differed from the mean won by the normal control group. BL OFC = bilateral orbito-frontal cortex; UL OFC = unilateral orbito-frontal cortex; Medial PFC = medial prefrontal cortex; DLPFC = dorsolateral prefrontal cortex. Attended/failed to attend = the patient attended/failed to attend to the crucial feedback during the reversal test, namely, the amount won or lost on each trial.

visual discrimination reversal test. However, in this previous study, the more diffuse nature of the lesions (with some produced by closed head injury, although nevertheless with clear damage to the ventral prefrontal cortex evident in structural MRI scans) precluded precise localization of this impairment to the orbito-frontal cortex. In the present study, by investigating patients with surgical and thus more circumscribed lesions of the frontal lobes, we were able to demonstrate that an impairment on reversal learning can occur following bilateral orbital excisions. These results are consistent with findings from nonhuman primates, which implicate the orbito-frontal cortex in reward and in stimulus–reward learning and reversal (Dias et al., 1996; Rolls, 1990, 1999b, 2000) as well as with the results from our neuroimaging study using the same probabilistic reversal task (O'Doherty, Kringelbach, et al., 2001), which produced bilateral activation in the orbito-frontal cortex in normal subjects. In the present study, although as a group the patients with bilateral orbital lesions were significantly impaired on this test, only those whose lesions included the areas activated in the fMRI study showed a marked impairment, whereas the two patients with bilateral lesions who were not significantly impaired (V.O. and R.R.) had small lesions restricted either to the region anterior or to the region posterior to the regions of activation (see Figure 3A).

In the fMRI study, it was found that the medial and the lateral regions of the orbito-frontal cortex were activated, respectively, by reward and by punishment, and indeed that the degree of activation in these two regions was correlated with the amount of ''money'' gained or lost (O'Doherty, Kringelbach, et al., 2001). In the present study, the patients with bilateral orbital lesions who were impaired on the test all had lesions that encompassed both the more medial and the more lateral regions described in the activation study, and for that reason, a difference in sensitivity to reward and to punishment would not be expected in this patient group. Consistent with this, no clear pattern of greater sensitivity to punishment or to reward was observed in those with bilateral lesions who were impaired. Further research with more discrete lesions within the orbito-frontal cortex may reveal more detailed parallels with the fMRI study, perhaps with sensitivity to reward being impaired by medial orbito-frontal cortex lesions and sensitivity to punishment by lateral orbito-frontal cortex lesions. However, since a recent fMRI study found that ventrolateral activation was associated with reversal learning independently of the effects of negative feedback (Cools, Clark, Owen, & Robbins, 2002), the medial–lateral distinction (within the orbital cortex) for reward and for punishment respectively seems less likely to hold up for this kind of test.

It is important that, in the present study, the group with unilateral orbital lesions on either side was unimpaired, indicating that it is necessary for the orbital damage to extend bilaterally before impairments are shown on visual discrimination reversal learning. This suggests that either the left or the right orbito-frontal cortex on its own can carry out the functions necessary for normal performance. It is very unlikely that lesion size per se can explain this effect, because some patients with unilateral orbital lesions who were unimpaired had a larger total area of prefrontal cortex excised than some of the patients with bilateral orbital lesions who were impaired. This was confirmed by the fact that there was no correlation between the total lesion size (amount of prefrontal cortex excised) and performance on the task.

The design of the probabilistic reversal task means that two interpretations of the impairment shown by patients with bilateral orbital/medial lesions can be excluded. First, the repeated touching of the previously rewarded stimulus cannot be explained in terms of simple motor perseveration (i.e., as a failure of inhibitory control of the arm/hand with which the patients reached out to touch the now-incorrect stimulus). Because a response was required and made on every trial, the continued selection of the old S+ must reflect instead the patient's difficulty, when reinforcement contingencies are reversed, in altering the previous association formed between the stimulus and reward or that between the stimulus and the response he should make to it. Although the present study does not distinguish between these two interpretations, the results are consistent with the hypothesis (Rolls, 1990, 1999b, 2000, 2002) that one important function of the orbito-frontal cortex consists in representing the reward and punishment value of stimuli, and in particular, since all patients succeeded in the early stages of the task (before the first reversal), in updating the associations between stimuli and reward or punishment value when these change.

Secondly, the clear instructions that explained that reversals would occur and that subjects were to alter their choice of stimulus accordingly exclude an interpretation in terms of lack of initiative. This might have explained some of the problems shown by patients in our earlier study (Rolls et al., 1994) in which no warning was given that the contingencies would change. The fact that clear explanation was given about this in the present study at least excludes an explanation in these terms. However, we cannot exclude the possibility that some of the patients with dorsolateral lesions who were impaired at the test and who failed to attend to the essential feedback had failed fully to understand the task requirements, and if so, this itself may be viewed as an aspect of executive dysfunction. Interestingly, among those who failed to attend to the amount won or lost on each trial, some did attend to this for some of the time, but did not appreciate its importance and started to attend to other sources of feedback instead (Appendix 2B).

The difficulty in selecting the appropriate feedback, or using it in a consistent manner, which was shown by

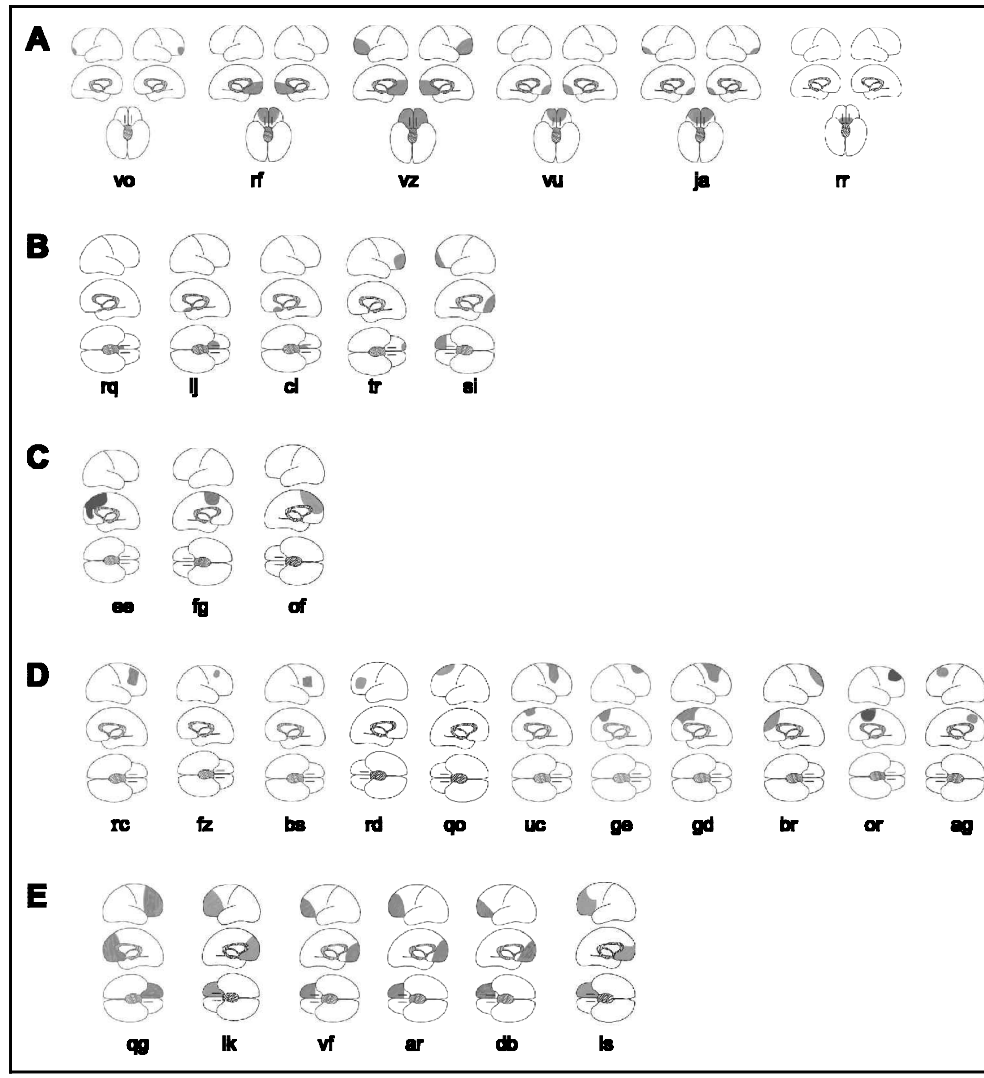

Figure 3. Brain maps showing lesion sites. Within each group with unilateral lesions, those with right-sided lesions are shown first. (A) Bilateral orbito-frontal cortex (or orbital plus medial prefrontal cortex). (B) Unilateral orbito-frontal (or orbital plus medial). (C) Unilateral medial prefrontal cortex. (D) Unilateral dorsolateral prefrontal (or dorsolateral plus medial prefrontal). (E) Unilateral orbito-frontal plus medial prefrontal plus dorsolateral prefrontal cortex.

almost half of the patients in the dorsolateral group, and which resulted in severely impaired performance on the reversal test, has parallels in work with nonhuman primates in which dorsolateral lesions impaired the ability to "monitor multiple stimuli" (Petrides & Pandya, 1999), as well as with neuroimaging studies of executive functions, such as those involved in dual-task performance (Szameitat, Torsten, Muller, & von Cramon, 2002). On the posttest questionnaire, the patients in this study reported that, instead of using the amount won or lost on each trial to guide their choice of pattern, they had attended instead to other nonessential sources of information, such as the cumulative total (as a number or as a bar chart), whether they had won or lost on each trial (ignoring the amount), or the colors or sounds that were associated with gains or losses. (We note that the self-report provided by these patients with circumscribed frontal lesions was sufficiently reliable that across the group the self-report about attention to the essential feedback was strongly related to the level of performance on the task. Indeed, there was a significant difference between the proportions of patients who were impaired who attended, 2/9, and who failed to attend, 7/8; Fisher's exact test, $p = .02$, two-tailed.) It is of interest that the dorsolateral group who failed to attend, or who failed to attend consistently, to the essential feedback during the main reversal phase of the test were able to appreciate which feedback to use in the far easier practice session of the test in which no reversals occurred. A similar effect of task difficulty in revealing impairments in patients with dorsolateral lesions has been reported by Morris et al. (2002), who found that the ability to employ the appropriate strategy in a test of planning ability, was not apparent in the early part of the test but became apparent only in the later, harder phase. Similarly, in Szameitat et al.'s (2002) study, dorsolateral prefrontal cortex activation was produced only when the task difficulty was increased.

In the present study, three of the four (R. C., F. Z., and B. S.) whose lesions were confined to the dorsolateral prefrontal region all failed to attend to the amount won/lost on each trial, suggesting that a dorsolateral

lesion may be sufficient to produce this cognitive/executive impairment. In those cases where the dorsolateral lesions extended down into the medial region, there are also parallels with imaging studies in which the more dorsal regions of anterior cingulate, described by the authors as its cognitive division, are activated in normal subjects during tasks involving stimulus–response selection in the face of competing streams of information (Bush, Luu, & Posner, 2000). The dorsolateral prefrontal cortex lesions did in fact extend into this ''cognitive division'' region of the anterior cingulate cortex in the other five patients with dorsolateral lesions (G.E., G.D., O.R., A.G., and L.K.) who also failed to attend to the appropriate feedback. However, since other patients with comparable lesions extending into these regions succeeded in attending to this feedback, this suggestion must remain tentative. It will be of interest to develop our reversal task in the future to simplify the feedback provided in the expectation that this will make the task performable without deficit by patients with dorsolateral prefrontal cortex lesions, but still sensitive to the effects of lesions to the orbito-frontal cortex due to the reversal component of the task.

The findings reported in this study suggest that the probabilistic reversal test makes demands on different functions attributed, respectively, to the dorsolateral and to the orbital regions of the prefrontal cortex (cf. Bechara et al., 1998, who reported a study in which lesions in these regions had dissociable effects on decision making and on working memory). Thus, in the present study, the need to determine which of multiple competing sources of feedback should be attended to, to maintain attention to this throughout the test, ignoring other visually salient but noncrucial types of feedback (flashing lights, total sum displayed, colored bar charts, etc.) are among the many functions that have been attributed to the dorsolateral prefrontal cortex (Szameitat et al., 2002; Duncan & Owen, 2000). For a subject who is capable of attending consistently to the appropriate feedback, it is also necessary to use the information about the current level of reward associated with each pattern to guide his choice on each trial, and as such, the test also makes demands on functions attributed to the orbito-frontal cortex (Rolls, 1990, 1999a, 1999b, 2000). The results of the present study show that it is possible to be severely impaired on this probabilistic visual discrimination reversal test for either of these two reasons, and in the context that performance on the test can be independently disrupted by the effects of dorsolateral and of orbital lesions, we have shown a dissociation between the effects of lesions in these two regions. The ability to attend to the crucial feedback can be disrupted by a dorsolateral lesion on either side, whereas the ability to use this information to choose advantageously can be supported by either orbital region on its own (after a unilateral orbital lesion), but is abolished by a bilateral orbital lesion. These results are consistent with the view that, whereas the dorsolateral frontal cortex forms part of a network of prefrontal regions recruited to solve diverse cognitive problems, especially those with an ''executive'' component (in this case, appreciating which feedback is relevant/important) (Bush et al., 2000; Duncan & Owen, 2000), the orbito-frontal cortex is involved in the representation of the changing reward value of stimuli (Rolls, 1990, 1999b, 2000, 2002) so that behavior can be modified accordingly. As such, the present study contributes both to our understanding of the different functions of these two regions of frontal lobe, and to how these regions need to cooperate on complex tasks, such as the one used here. A similar argument is made by Manes et al. (2002) who showed how the ventral and dorsal aspects of the prefrontal cortex must interact in the maintenance of rational and ''nonrisky'' decision-making tasks.

One of the underlying hypotheses of the research described here is that some of the reasons for the behavioral and emotional changes that follow damage to the prefrontal cortex may be related to deficits in decoding the reward and punishment value of stimuli and using the results of the decoding to modify behavior and the ongoing emotional or mood state (Rolls, 1990, 1999a, 1999b; Rolls et al., 1994). In this context, it is of interest that the patients with bilateral orbito-frontal cortex damage who were impaired at the visual discrimination reversal task (see Table 1) had high scores on parts of a social behavior questionnaire, in which the patients were rated on behaviors, such as the recognition of emotion in others (e.g., their sad, angry, or disgusted mood); in interpersonal relationships (such as not caring what others think, and not being close to the family); in sociability (is not sociable and has difficulty making or maintaining close relationships); emotional empathy (e.g., when others are happy is not happy for them); public behavior (is uncooperative); antisocial behavior (is critical of and impatient with others); and impulsivity (does things without thinking) (Hornak, Bramham, Rolls, Morris, O'Doherty, & Polkey, 2003), all of which could reflect less behavioral sensitivity to different types of punishment and reward. Further, in a subjective emotional change questionnaire, in which the patients reported on any changes in the intensity and/or frequency of their own experience of emotions since surgery, the bilateral orbito-frontal cortex patients with deficits in the visual discrimination reversal task reported a number of changes, including changes in sadness, anger, fear, and happiness, and were as a group significantly different in these respects from the dorsolateral prefrontal cortex lesion group included in that comparison (Hornak et al., 2003). However, it must also be pointed out that the patients with unilateral orbito-frontal/anterior cingulate lesions (the BA 9/ACC and BA 9/ACC + orbital groups of that article) also reported marked emotional changes and they were also less well-adjusted socially than the dorsolateral group, but had no measured deficit in reversal. The

involvement of the orbito-frontal cortex in reversal, as revealed by the performance of the patients with bilateral orbito-frontal cortex damage described in this article, as well as by our fMRI (O'Doherty, Kringelbach, et al., 2001), may mean that even subtle changes in the sensitivity to reward and nonreward in patients with unilateral lesions that include orbital regions, as well as in other medial regions closely connected with the orbito-frontal cortex, such as the anterior cingulate cortex, may be associated with deficits in what may be sensitive indicators of this type of function, such as subjective emotional experience and social behavior.

METHODS

Subjects

Patient Groups

Thirty-one patients were included in the study. They were under the care of the Department of Neurosurgery, King's College Hospital, London. Informed consent was obtained, and the study was approved by the Psychiatry Ethical Committee (Study 157/00). Exclusion criteria included damage outside the prefrontal cortex, alcohol- or drug-dependence, and a full-scale IQ below a cut-off of 80. The IQ for controls and for patients was estimated using the NART (Nelson, 1982). For the control group, mean IQ was 112.2 (*SD* 9.7); the patients IQs are shown in Table 3b, along with other clinical details. In addition to IQ, information about the educational level and occupation for the normal control group as well as for the different patient groups is shown in Appendix 1. These tables show that the patient groups were at least as well educated and had achieved a level of occupation at least as good as that of the normal control group.

Categorization of Lesions

Table 3a and Figure 3 show the lesion sites for each patient. The brain maps are based on the surgeons' drawings showing by direct visual observation which parts of the brain were removed. Only three patients (R.F., J.A., and Q.G.) had radiation therapy, and white matter changes associated with this cannot be excluded in these three patients. Any epilepsy patient in this study had focal epilepsy, and some were taking anticonvulsants at the time of testing as shown in Table 3b. The etiologies that are shown in Table 3b can be summarized as follows: 13 patients had suffered from meningioma, 11 from epileptic focus, 2 from focal head injury, and 1 from each of the following: astrocytoma, oligodendroglioma, cavernoma, malignant ependymoma, and anterior communicating artery aneurysm with subarachnoid hemorrage.

A method of categorization was used in which the patients were classified according to the prefrontal sectors of functional significance into which the lesions encroached (Rowe, Bullock, Polkey, & Morris, 2001; Rowe, Owen, Johnsruude, & Passingham, 2001). These sectors were defined anatomically as orbital (Brodmann's areas 10, 11, 12, and 25), medial (Brodmann's areas 8, 9, and 10), and dorsolateral (Brodmann's areas 9 and 46).

Table 3a. Lesion Sites: Classification According to the Main Subdivisions of Prefrontal Cortex

Patient	*Side*	*Orbital BA 10, 11, 12, 25*	*Medial BA 8, 9, 10*	*Dorsolatateral BA 9, 46*
V.O.	Bilat	(++) Fr. pole		
R.F.	Bilat	++	++	
V.Z.	Bilat	++	++	
V.U.	Bilat	++	++	
J.A.	Bilat	++	++	
R.R.	Bilat	++		
R.Q.	Right	+		
L.J.	Right	+		
C.L.	Right	+		
T.R.	Right	+		
S.I.	Left	+	+	
E.E.	Right		+	
F.G.	Left		+	
O.F.	Left		+	
R.C.	Right			+
F.Z.	Right			+
B.S.	Right			+
R.D.	Left			+
Q.O.	Left			+
U.C.	Right		+	+
G.E.	Right		+	+
G.D.	Right		+	+
B.R.	Right		+	+
O.R.	Right		+	+
A.G.	Left		+	+
Q.G.	Right	+	+	+
L.K.	Left	+	+	+
V.F.	Left	+	+	+
A.R.	Left	+	+	+
D.B.	Left	+	+	+
L.S.	Left	+	+	+

BA = Brodmann's area; Bilat = bilateral; + unilateral lesion in this region; ++ Bilateral lesions in this region.

Table 3b. Clinical Information About the Patients

Patient	Sex	Age	IQ (NART)[a]	Years Since Surgery	Etiology of Lesion
Bilateral orbital/medial					
V.O.[b]	M	48	124	0	Contusions: focal head injury (closed)
R.F.	M	48	116	4	Olfactory groove meningioma
V.Z.	M	72	102	18	Olfactory groove meningioma
V.U.	F	61	94	1	Meningioma
J.A.	F	60	128	6	Meningioma
R.R.	F	55		2	Subfrontal meningioma
Unilateral orbital/medial					
R.Q.	F	51	106	7	anterior communicating artery aneurysm and subarachnoid hemorrage
L.J.	F	41	118	1	Suprasellar meningioma
C.L.	M	70	126	7	Frontal planum sphenoidale meningioma
T.R.	F	52	106	5	Meningioma
S.I.	M	31	112	13	Hematoma: focal head injury (closed)
Unilateral medial					
E.E.	M	59	111	2	Meningioma
F.G.	M	57	122	4	Meningioma
O.F.	M	42	111	2	Epilepsy
Dorsolateral					
R.C.	M	31	81	14	Epilepsy
F.Z.	M	20	105	1	Cavernoma
B.S[b]	F	31	102	12	Epilepsy
R.D.	F	54	118	3	Meningioma
Q.O.	M	25	106	2	Oligodendroglioma
Dorsolateral/medial					
U.C.[b]	M	32	106	10	Epilepsy
G.E.[b]	F	25	101	13	Oligodendroglioma
G.D.	F	45	123	20	Epilepsy
B.R.[b]	F	19	97	4	Epilepsy
O.R.[b]	F	32	98	5	Epilepsy
A.G.[b]	F	34	98	0	Malignant ependymoma
Dorsolateral/medial/orbital					
Q.G.[b]	M	55	114	0	Meningioma
L.K.	M	30	112	10	Epilepsy
V.F.	M	32	116	12	Epilepsy
A.R.[b]	F	37	112	22	Epilepsy
D.B.	M	63	126	1	Astrocytoma
L.S.	M	32	120	10	Epilepsy

M = male; F = female.

[a]From Nelson (1982).

[b]The patient was taking medication at the time of testing (anticonvulsants).

This categorization produced the five groups of patients shown in Figure 3, and also shown in Tables 3a and 3b. (In Table 3b, the patients shown as dorsolateral and as dorsolateral/medial are both included in D in Figure 3.) As shown in Table 3a, those with bilateral lesions had orbital or a combination of orbital and medial lesions. There were no patients with bilateral dorsolateral lesions. Among patients with unilateral lesions, some had lesions confined to the orbital or medial or dorsolateral region, and others had lesions in two or three of these regions. The five groups of patients shown in Figure 3 (A–E) correspond to those used in related research on the same patients (Hornak et al., 2003). For the purposes of the analyses described in this article, all patients with lesions in the dorsolateral prefrontal cortex (Groups D and E in Figure 3) were treated as a single group of patients with damage that included dorsolateral prefrontal cortex damage (the dorsolateral prefrontal cortex group).

Control Group

Twenty-five normal subjects took part, matched for sex, age, level of education, occupational category, and IQ measured using NART (Nelson, 1982).

Reversal Test

The task consisted of a visual (object) discrimination learning test, in which the same two previously unfamiliar fractal patterns were used throughout, as shown in Figure 4. On each trial, these appeared together, one above the other, on a touch screen, and the subjects task was asked to choose one by touching it. The

position of each pattern (above or below) was randomized using a pseudorandom order.

Reward Contingencies

The rewards and losses, which could be obtained for a choice of a stimulus, were distributed probabilistically, varying in both frequency and magnitude. The frequency ratio of rewards to losses for the "good" stimulus (S+) was 70:30, whereas for the "bad" stimulus (S−) it was 40:60. The magnitude of the rewards varied according to a uniformly distributed random sequence, as follows: for the S+, the rewards ranged from £80 to £250 and the losses ranged from £10 to £60. For the S−, the rewards ranged from £30 to £65 and the losses ranged from £250 to £600. Feedback was given on each trial, as shown in Figure 4.

Feedback

There were two main types of feedback that were important in allowing the patient to monitor his success on the task. These were;

Amount won or lost on each trial. This was signaled prominently by a message in flashing colors that was superimposed upon the stimulus for 3 sec. For a win, the message was "WELL DONE! YOU HAVE WON £…", and for a loss, the message was "SORRY! YOU HAVE LOST £…."

Cumulative total gained/lost so far. This appeared numerically and as a bar chart. Numerically, it appeared as "TOTAL: £…" in the middle of the screen, between the two stimuli, and was updated after each trial. The sum could be preceded by a plus or minus sign and started at zero at the beginning of the test. A bar chart representing the total gained or lost, with a horizontal line indicating zero, was also shown to the right of the screen, moving above the horizontal for positive values and below it for negative.

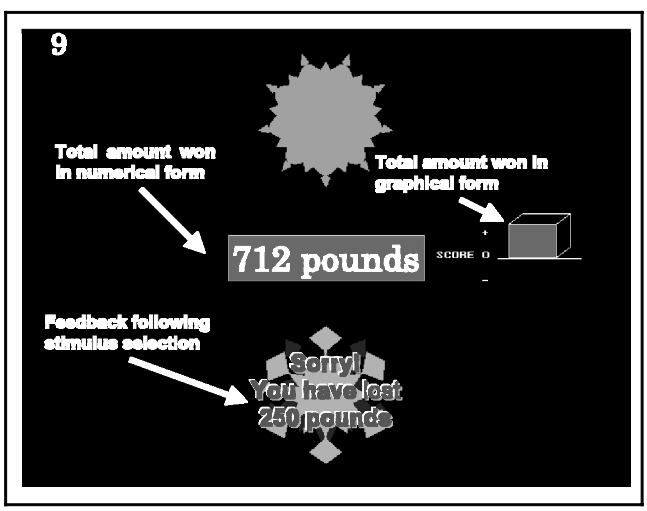

Figure 4. Example of a trial from the Reversal test showing the stimuli and the different sources of feedback displayed on the touchscreen after each choice.

Additionally, two further sources of feedback were included to increase the salience of whether the subject had just won or lost by the choice he/she had just made. These were;

Color. For messages and for the bar chart, blue was used to signal gains/positive values, whereas red signaled loss/negative values.

Sound. Gains were accompanied by an ascending, and losses by a descending, frequency-modulated tone, and the duration of the sound was directly proportional to the amount won or lost (ranging from 100 msec to 1.5 sec).

The essential feedback from the point of view of succeeding on the task was the amount won or lost on each trial. The total accumulated also provided information about how well the subject was doing overall, but a subject who relied only on the changing total to guide his/her responses would need to calculate how much the pattern he had just selected had won or lost him/her. To accomplish this, he would need to compare the current total with that which had been displayed prior to his last choice, and this would require him to remember this number and to engage in mental arithmetic (addition or subtraction) before he could decide whether he should stay with the pattern he had just chosen or switch to the other one.

The additional feedback (bar chart for total, colors, and sounds) was included to make the gaining or losing of money more salient for the subject.

Practice

Before the reversal test proper, the subjects performed a version of the task (acquisition only) in which there was no reversal. This allowed subjects to familiarize themselves with the task demands and to appreciate the probabilistic nature of the reward contingencies. Subjects practiced until they had reached a criterion of selecting the S+ on 16 out of the 18 preceding responses. For this version of the test, the ratios of rewards and losses on the S+ and on the S− were the same as for the main reversal test, but the amounts of reward and loss were different. For the S+, the rewards ranged from £60 to £200 and the losses from £10 to £50. For the S−, the rewards ranged from £10 to £100 and the losses from £70 to £300.

Main Reversal Test

For the main part of the test, which followed the practice, two new fractal patterns were used. All subjects started

with £0. After a criterion of nine selections of the S+ out of the preceding 10 trials was reached, the first reversal began. This was achieved by gradually reversing, in incremental steps over a period of 10 trials, the reward contingencies of the two stimuli. The probabilities shifted in a regular stepwise fashion over the 10 trials. The test stopped after 100 trials for all subjects. The subject's comments were tape-recorded during the test.

Instructions to the Subject

Practice. The subjects were informed that each pattern, when touched, could either give or take away varying amounts of imaginary "money" and that this would be displayed on the screen. The different types of feedback were described and demonstrated on the screen. It was explained that one pattern gave more than it took overall (the "good" pattern) and that one took more than it gave overall (the "bad" pattern). The subjects were told that if they kept choosing the good pattern they would earn lots of "money," whereas if they kept choosing the "bad" pattern, they would lose a lot of money, and would "go more and more into the red" (into minus figures). Their task was to determine, by trial and error, which pattern it was more profitable to choose (by touching it) and to stick with it.

Reversal test. The subjects were told that this part of the test would start in the same way as the practice, but that once they had found out which was the good pattern and had touched it consistently a number of times, it would gradually become "bad," and that the bad pattern would become gradually become "good." The instructions were:

> Your aim is to adjust your choices accordingly—to start choosing the pattern that had been bad at the beginning and to avoid choosing the pattern which had started off being good. Once you have successfully switched your choice of pattern and have chosen it consistently a certain number of times, there will be a second reversal, back to how things were at the beginning, and later a third reversal, and so on. Your aim is to win as much money as possible by keeping track of which pattern is currently the good pattern and choosing it consistently until you think it is changing and becoming the bad pattern.

Posttest Questionnaire

Directly after the subject had completed 100 trials, the subject was asked the open-ended question: "What information on the screen did you find most useful in keeping track of which pattern was currently the good one?" After the subjects had answered this question, they were then probed about each of the other sources of feedback in turn. Their responses were tape-recorded. The testing and the posttest questionnaire were carried out before details of the patients' lesions were available, and in this sense the testing was performed blind.

Acknowledgments

This research was supported by the Medical Research Council Interdisciplinary Research Centre in Cognitive Neuroscience.

Reprint requests should be sent to E. T. Rolls, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK, or via e-mail: edmund.rolls@psy.ox.ac.uk.

REFERENCES

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron, 32,* 537–551.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50,* 7–15.

Bechara, A., Damasio H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience, 18,* 428–437.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience, 2,* 382–387.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy of monetary gains and losses. *Neuron, 30,* 619–639.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215–222.

Butter, C. M. (1969). Perseveration in extinction and in discrimination reversal tasks following selective prefrontal ablations in *Macaca mulatta*. *Physiology and Behavior, 4,* 163–171.

Cools, R., Clark, L., Owen, A., & Robbins, T. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience, 22,* 4563–4567.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380,* 69–72.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience, 23,* 457–483.

Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: Evidence form human neuroimaging studies. *Cerebral Cortex, 10,* 308–317.

Elliot, R., Frith, C. D., & Dolan, R. J. (1997). Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia, 35,* 1395–1404.

Francis, S., Rolls, E. T., Bowtell, R., McGlone, F., O'Doherty, J., Browning, A., Clare, S., & Smith, E. (1999). The representation of pleasant touch in the brain and its relationship with taste and olfactory areas. *NeuroReport, 10,* 453–459.

Frey, P., Kostopoulos, P., & Petrides, M. (2000). Orbitofrontal involvement in the processing of unpleasant auditory information. *European Journal of Neuroscience, 12,* 3709–3712.

Goldman-Rakic, P. S. (1996). Regional and cellular fractionation of working memory. *Proceedings of the National Academy of Sciences, U.S.A., 93,* 13473–13480.

Gottfried, J. A., Deichmann, R., Winston, J. S., & Dolan, R. J. (2002). Functional heterogeneity in human olfactory cortex: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 22,* 10819–10828.

Hornak, J., Bramham, J., Rolls, E. T., Morris, R. G., O'Doherty, J., Bullock, P. R., & Polkey, C. E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain, 126,* 1691–1712.

Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia, 34,* 247–261.

Iversen, S. D., & Mishkin, M. (1970). Perseverative interference in monkey following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research, 11,* 376–386.

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

Manes, F., Sahakian, B. B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain, 125,* 624–639.

Meunier, M., Bachevalier, J., & Mishkin, M. (1997). Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. *Neuropsychologia, 35,* 999–1015.

Morris, R. G., Kotisa, M., Brooks, B., Rose, F. D., Bullock, P., & Polkay, C. E. (2002). *Virtual reality investigations of planning ability following focal prefrontal cortical lesions.* Paper presented at the 19th Annual British Psychological Society, Cognitive Psychology Section Conference, Kent, UK.

Nelson, E. H. (1982). *National Adult Reading Test (NART): Test Manual.* Windsor: NFER-Nelson.

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

O'Doherty, J., Rolls, E. T., Francis, S., Bowtell, R., McGlone, F., Kobal, G., Renner, B., & Ahne, G. (2000). Sensory-specific satiety related olfactory activation of the human orbitofrontal cortex. *NeuroReport, 11,* 893–897.

O'Doherty, J., Rolls, E. T., Francis, S., McGlone, F., & Bowtell, R. (2001). Representation of pleasant and aversive taste in the human brain. *Journal of Neurophysiology, 85,* 1315–1321.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41,* 147–155.

Owen, A. M. (1997). Cognitive planning in humans: Neuropsychological, neuroanatomical and neuropharmacological perspectives. *Progress in Neurobiology, 53,* 431–450.

Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex : Comparative cytoarchitectonic analysis in the human and macaque brain and cortico-cortical connection patterns. *European Journal of Neuroscience, 11,* 1011–1036.

Robbins, T. W. (1998). Dissociating executive functions of the prefrontal cortex. In A. C. Roberts (Ed.), *Functions of the prefrontal cortex* (pp. 117–130). Oxford: Oxford University Press.

Roberts, A. C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex, 10,* 252–262.

Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., & Robbins, T. W. (2000). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *Journal of Cognitive Neuroscience, 12,* 142–162.

Rolls, E. T. (1990). A theory of emotion, and its application to understanding the neural basis of emotion. *Cognition and Emotion, 4,* 161–190.

Rolls, E. T. (1999a). The functions of the orbitofrontal cortex. *Neurocase, 5,* 301–312.

Rolls, E. T. (1999b). *The brain and emotion.* Oxford: Oxford University Press.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10,* 284–294.

Rolls, E. T. (2002). The functions of the orbitofrontal cortex. In D. T. Stuss & R. T. Knight (Eds.) *Principles of frontal lobe function* (chap. 23, pp. 354–375). New York: Oxford University Press.

Rolls, E. T., Critchley, H., Mason, R., & Wakeman, E. A. (1996). Orbitofrontal cortex neurons: Role in olfactory and visual association learning. *Journal of Neurophysiology, 75,* 1970–1981.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery and Psychiatry, 57,* 1518–1524.

Rolls, E. T., Kringelbach, M. L., De Araujo, I. E. T. (2003). Different representations of pleasant and unpleasant odors in the human brain. *European Journal of Neuroscience, 18,* 695–703.

Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., & McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cerebral Cortex, 13,* 308–317.

Rolls, E. T., Sienkiewicz, Z. J., & Yaxley, S. (1989). Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *European Journal of Neuroscience, 1,* 53–60.

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). "Theory of mind" impairments and their relationships to executive functioning following frontal lobe excisions. *Brain, 124,* 101–117.

Rowe, A. D., Owen, A. M. Johnsruude, I. S., & Passingham, R. E. (2001). Imaging the mental components of a planning task. *Neuropsychologia, 39,* 315–327.

Small, D. M., Zald, D. H., Jones Gotman, M., Zatorre, R. J., Pardo, J. V., Frey, S., & Petrides, M. (1999). Human cortical gustatory areas: A review of functional neuroimaging data. *NeuroReport, 10,* 7–14.

Szameitat, A. J., Torsten, S., Muller, K., & von Cramon (2002). Localization of executive function in dual-task performance with fMRI. *Journal of Cognitive Neuroscience, 14,* 1184–1199.

Thorpe, S. J., Rolls, E. T., & Maddison, S. (1983). Neuronal activity in the orbitofrontal cortex of the behaving monkey. *Experimental Brain Research, 49,* 93–115.

Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex–basal ganglia system in primates. *Critical Reviews in Neurobiology, 10,* 317–356.

Zald, D. H., Lee, J. T., Fluegel, K. W., & Pardo, J. V. (1998). Aversive gustatory stimulation activates limbic circuits in humans. *Brain, 121,* 1143–1154.

Zald, D. H., & Pardo, J. V. (1997). Emotion, olfaction and the human amygdala: Amygdala activation during aversive olfactory stimulation. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 4119–4124.

APPENDICES

APPENDIX 1. EDUCATIONAL AND OCCUPATIONAL LEVEL

Patients are grouped as in Tables 3a, 3b and Figure 3.

Educational Level	*Normal Control (n = 25)*	*BL OFC (n = 6)*	*UL OFC (n = 5)*	*MED (n = 3)*	*DL/MED (n = 10)*	*DL/MED/ORB (n = 6)*
Left school (%)	56	60	40	33	64	33
Degree (%)	40	0	40	33	18	33
Postgraduate (%)	4	40	20	33	18	33

Left school = left school without going on to higher education.

Occupational Level	*Normal Control*	*BL OFC*	*UL OFC*	*MED*	*DL/MED*	*DL/MED/ORB*
Semiskilled (%)	8	0	20	0	27	0
Skilled (%)	40	0	20	0	9	17
Managerial (%)	36	60	20	67	46	33
Professional (%)	16	40	40	33	18	50

BL OFC = bilateral orbital/medial; UL OFC = unilateral orbital/medial; MED = medial; DL/MED = dorsolateral/medial frontal cortex; DL/MED/ORB = dorsolateral/medial/orbital.

APPENDIX 2A. POSTTEST QUESTIONNAIRE

Patient	*Earnings on Each Trial*	*Cumulative Total: Numerical*	*Cumulative Total: Bar Chart*	*Sound*	*Color*
Without dorsolateral lesion					
Bilateral orbital/medial					
R.F.	Y	Y	N	Y	N
V.Z.	Y	(Y)	N	(Y)	N
V.U.	Y	Y	Y	Y	Y
J.A.	Y	Y	N	(Y)	Y
R.R.	Y	Y	(Y)	(Y)	N
V.O.	Y	N	N	(Y)	(Y)
Unilateral orbital/medial					
R.Q.	Y	Y	N	Y	N
L.J.	Y	Y	(Y)	(Y)	Y
C.L.	Y	(Y)	N	N	N
T.R.	Y	N	N	N	N
S.I.	Y	(Y)	N	N	N
Unilateral medial					
O.F.	Y	(Y)	N	N	N
F.G.	Y	(Y)	N	(Y)	N
E.E.	Y	N	N	N	N

APPENDIX 2A. (*continued*)

Patient	*Earnings on Each Trial*	*Cumulative Total: Numerical*	*Cumulative Total: Bar Chart*	*Sound*	*Color*
With dorsolateral lesion (attended)					
U.C.	Y	N	N	N	N
Q.O.	Y	Y	(Y)	N	N
B.R.	Y	(Y)	N	N	(Y)
Q.G.	Y	(Y)	N	(Y)	N
V.F.	Y	N	N	N	N
A.R.	Y	(Y)	Y	Y	N
D.B.	Y	Y	Y	(Y)	Y
R.C.	N	Y	(Y)	(Y)	(Y)
F.Z.	N[a]	Y	(Y)	Y	Y
B.S.	N	(Y)	(Y)	(Y)	(Y)
G.E.	N	Y	Y	N	Y
G.D.	N[b]	Y	Y	Y	Y
O.R.	N	Y	Y	N	N
A.G.	N[c]	Y	(Y)	Y	Y
L.K.	N	Y	N	Y	Y

''Attended/failed to attend'': These patients attended or failed to attend to the essential feedback, namely, the amount won or lost on each trial. (Y): The patient noticed these sources of feedback but they did not consider them important. They paid attention some of the time/kept an eye on them, but did not use them to guide their choice of pattern on the next trial.

[a]F.Z. attended to the amount earned on each trial at the beginning. ''Then other things seemed more important'' (see Appendix 2B).

[b]G.D. attended to the amount earned on each trial to begin with ''but later other things seemed a better guide.''

[c]A.G. only noticed whether she had won or lost on each trial, but ''paid very little attention'' to the amount.

APPENDIX 2B. NONESSENTIAL FEEDBACK USED BY PATIENTS WITH DORSOLATERAL LESIONS

Some patients with dorsolateral lesions, when questioned about the feedback they had used to determine their choice of pattern during the reversal task, said that they had noticed only whether the message on the screen indicated whether they had won or lost on each trial, but had not paid attention to the actual amount. Many of these reported that their choice of pattern had been determine by their total sum (which was displayed both as a number in the center of the screen or as a bar chart). Some said they had been guided by the flashing of messages or the colors or the sounds that accompanied gains and losses. Others said they had attended to the actual amount won or lost on each trial some of the time, but had not done so consistently. For example F. Z., who was severely impaired on the task (performing two standard deviations below the mean for the normal control group), when questioned about which feedback he had used task, said, ''I kept looking at the bar chart for some reason.'' When questioned about the amount won or lost on each trial, he said, ''At first I did notice it, then I noticed the flashing and thought it meant I was losing. Then I noticed it flashed for winning as well . . . so towards the end I noticed the amount on each go.'' Importantly, his performance mirrored this: He did well at first, then lost a large amount of money, then began doing better toward the end. Similarly, G.D. commented: ''I did pay attention to the amount on each trial at first, then other things seemed more important.''