ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

NeuroImage

NeuroImage 20 (2003) 1371–1383

www.elsevier.com/locate/ynimg

Neural correlates of rapid reversal learning in a simple model of human social interaction

Morten L. Kringelbach[a,b,*] and Edmund T. Rolls[a]

[a] *Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK*
[b] *Oxford Centre for Functional Magnetic Resonance Imaging (FMRIB), John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK*

Received 22 May 2003; revised 24 June 2003; accepted 26 June 2003

Abstract

Humans and other primates spend much of their time engaged in social interactions where a crucial ability is to decode face expressions and act accordingly. This rapid reversal learning has been proposed to be important in the relative evolutionary success of primates. Here we provide the first neuroimaging evidence that the ability to change behaviour based on face expression in a model of social interactions is not reflected in the activity in the fusiform face area, but is specifically correlated with activity in the orbitofrontal and anterior cingulate/paracingulate cortices. These brain regions are particularly involved in reversal learning, such that the activations described occurred specifically at the time of reversal, and were also found when different face expressions other than angry were used to cue reversal. The evidence that the orbitofrontal and anterior cingulate/paracingulate cortices are specifically activated at the time of reversal is important for understanding changes in affect and emotional processing in patients with lesions to these brain regions.

Keywords: Orbitofrontal cortex; Mentalizing; Functional magnetic resonance imaging; Reversal learning; Face expression

Introduction

In humans and other primates face expressions act as important social cues to regulate behaviour (Darwin, 1872; Ekman and Friesen, 1971; deWaal, 1997). Much is known about the neural correlates of the decoding of face expressions from neurophysiological studies in nonhuman primates (Bruce et al., 1981; Desimone and Gross, 1979; Hasselmo et al., 1989a, 1989b; Perrett et al., 1982; Rolls and Deco, 2002), and from human lesion (Adolphs et al., 1994; Bodamer, 1947; Sergent and Villemure, 1989) and imaging studies (Haxby et al., 1991, 1994; Köhler et al., 1995), but almost nothing is known about the neural correlates of how face expressions govern human social behaviour. On the basis of human lesion studies it has been proposed that the orbitofrontal cortex is crucial for controlling affective responses to salient emotional stimuli (Damasio, 1994; Davidson et al., 2000; Hornak et al., 1996, 2003a; Rolls, 1999; Rolls et al., 1994). This affective behaviour has been called "Machiavellian" or social intelligence (deWaal, 1997; Whiten and Byrne, 1997) and has been proposed to be distinct from other kinds of intelligence, which fits well with evidence from patients with damage to the orbitofrontal cortex who have been shown to have severe problems in social interactions despite performing well on conventional intelligence tests (Blair and Cipolotti, 2000; Eslinger and Damasio, 1985).

At the heart of social intelligence is the ability to detect subtle changes in communication and act upon these changes rapidly as they occur. Changes in social behaviour based on face expression come so naturally to humans and are in place so early in child development that some might argue that this functionality is essentially innate. However, human social behaviour is sufficiently flexible that we can easily learn to adapt our behaviour to most face expressions. While an angry face expression can be argued to be the natural signal in most situations for signalling that a change should occur in the current behaviour, we can easily learn that other face expressions also can signal a change in

* Corresponding author. Fax: +44-1865-310447.
E-mail address: Morten.Kringelbach@psy.ox.ac.uk (M.L. Kringelbach).

1053-8119/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S1053-8119(03)00393-8

behaviour. For example, neutral face expressions do not normally signal that behaviour should change, but it is easy to think of social contexts where a neutral face expression does indeed signal that our current behaviour is inappropriate and should change. This flexibility in social behaviour is an important aspect of much human behaviour, as it allows us to learn to adapt to the different behaviours of both individuals and groups.

The aim of the functional magnetic resonance imaging (fMRI) investigation described here was to identify the brain regions involved in changing behaviour in a simple model of human social and emotional interactions. To do this, we devised a "reversal" task, which captures some of the essence of learning to alter social behaviour using changes in face expression (see Fig. 1). Two versions of the task used different face expressions to cue the subject that a change or "reversal" of behaviour was needed. The overall goal of both tasks is for the subject to keep track of the mood of two people presented in a pair and as much as possible to select the "happy" person (who will then smile). Over time the person with the "happy" mood (who will smile when selected), changes the mood to "angry," indicated in the task by this person no longer smiling when selected, but instead changing to a face expression that signals that this person should no longer be selected. In the main reversal task the face expression used to cue reversal is the most natural face expression, namely, an angry face expression, while in the second, control, version of the reversal task, a neutral face expression is used instead to cue reversal. By using two different reversal tasks where different face expressions signal that behaviour must change, we were able to determine which brain areas are specific to general reversal learning rather than just to reversal following a particular face expression such as anger. To further rule out the effects of using angry face expressions to cue reversal in the brain regions related to reversal, we also used a modified version of the task where both angry and happy faces were presented to the subjects, but there was no need to change behaviour.

Materials and methods

Subjects

Nine healthy, right-handed subjects participated in this study (six female, three male). Subjects gave informed consent in accordance with full ethical approval by the Central Oxford Research Ethics Committee (COREC).

Experimental design

In the main reversal task two people with neutral face expressions are presented to the subject as shown in Fig. 1. The beginning of each trial is synchronized with a new scan. The subject has to select one of the people by pressing the corresponding button, and the person will then either smile or show an angry face expression depending on the current mood of the person. The task for the subject is to keep track of the mood of each person and choose the "happy" person as much as possible. Over time (after between four and eight correct trials) this will change so that the "happy" person becomes "angry" and vice versa, and the subject has to learn to adapt her choices accordingly. Randomly intermixed trials with either two men or two women as the stimuli were used to control for possible gender and identification effects. After the subject has pressed the button, the selected person's face is updated with the currently associated face expression for 3000 ms, and then in the intertrial interval a fixation cross is presented for 16 s (plus the time for the scanner to be ready to start a new scan, which is always less than 3 s, given that the TR is 3 s). Trials therefore typically lasted either 21,000 or 24,000 ms depending on the timing of the subject's response. Order effects were controlled for by the cyclical nature of the mood changes of the different men and women. One particular person in a pair will stay "happy" until chosen consistently for between four and eight trials, at which point this person becomes "angry" and stays "angry" until the other person in a pair (who is now "happy") has been consistently chosen for between four and eight trials, at which time the original person becomes "happy" again, and so on. [We note that in the task the subjects did not know when a reversal trial would occur, and consistently the choice times were not significantly different on reversal versus acquisition trials (1618 $\pm$ 127 ms mean $\pm$ SEM, vs 1630 $\pm$ 54 ms). Even on the trial that followed a reversal trial, and the subjects were selecting a different face to previously, there was only a small, not significant ($P < 0.05$), increase in response time to 1725 $\pm$ 115 ms.]

Two control tasks were devised as variations of the main reversal task to control for the face expression that was being used to cue reversal, and to rule out changes specific to using the most natural face expression, namely, angry, to cue reversal. The angry face is arguably the most commonly used face expression in human social interactions to convey that behaviour should change in others.

The control reversal task shares the reversal component with the main reversal task, but controls for face expression by using different face expressions to cue reversal. The control reversal task was modelled exactly like the main reversal task, but used different face expressions as stimuli and reinforcers. In this control reversal task the face expression used for both people during the stimulus presentation period of Fig. 1 was no longer neutral, but was instead angry. The subject still had to select the "happy" person as much as possible. The "happy" person is still the person who smiles when selected. However, in this control reversal task a required change in behaviour is signalled by a change to an indifferent (neutral) face expression (rather than to an angry face expression as in the standard reversal task). This change from an angry to a neutral face expression as a marker for a required change in behaviour is less natural

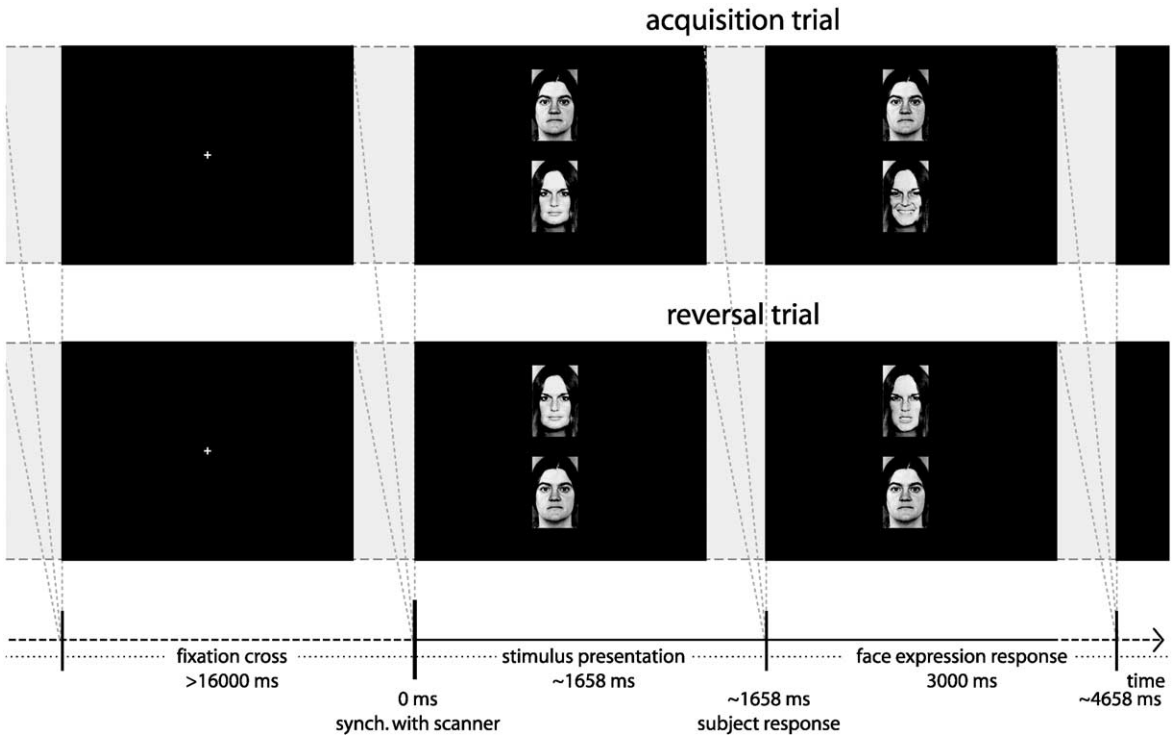


Fig. 1. Experimental design. Three variations on a reversal task were used in this study. The figure shows the main reversal task with a timeline of the events. The trial starts synchronised with the scanner, and two people with neutral face expressions are presented to the subject. The subject has to select one of the people by pressing the corresponding button, and the person will then either smile or show an angry face expression for 3000 ms depending on the current mood of the person. The task for the subject is to keep track of the mood of each person and choose the "happy" person as much as possible (top). Over time (after between four and eight correct trials) this will change so that the "happy" person becomes "angry" and vice versa, and the subject has to learn to adapt her choices accordingly (bottom). Randomly intermixed trials with either two men or two women were used to control for possible gender and identification effects, and a fixation cross was presented between trials for at least 16000 ms.

than the change from a neutral to an angry face expression in the main reversal task, but the subjects could still learn the task quite easily.

The face expression control task was designed to control for the effects of using an angry face expression as a cue to change behaviour. This control task is a very simple variation of the reversal task which does not have the reversal component. Two people with neutral expressions are still presented to the subject, who is instructed to find the person whose face expression changes randomly to the same face expressions as in the main reversal task, and to then keep choosing this person. In contrast to the main reversal task, a change to an angry face expression does not signal that a change in behaviour is needed and thus the face expression control task does not have a reversal component. Even if the subject does change behaviour to choose the other person, this person's face remains unresponsive (neutral).

Functional magnetic resonance imaging

Images were acquired with a 3.0-T Varian/Siemens whole-body scanner at FMRIB, Oxford. Twenty-one T2*-weighted EPI slices were acquired every 3 s (TR = 3). We used the techniques that we have developed over a number of years to carefully select the imaging parameters to minimise susceptibility and distortion artefact in the orbitofrontal cortex as described in detail by Wilson et al. (2002). Firstly, the data were acquired in a coronal rather than axial slicing direction, as this aligned the slices to be perpendicular to the predominant direction of the intrinsic susceptibility induced field gradients, and helps to minimise through-plane dephasing. Secondly, the voxel resolution was minimised by using 3-mm in-plane resolution and a 5-mm slice thickness which results in less phase cancellation than would be produced by lower voxel resolutions. Thirdly, a relatively low TE of 25 ms was selected to decrease the signal dropout, as less phase dispersion is created across the voxels. Fourthly, each subject was individually shimmed using both linear and second-order shimming to minimise static field inhomogeneities in the orbitofrontal cortex. Finally, geometric distortion was minimised by using a specialist head insert gradient coil (Magnex SGRAD III) with a relatively high gradient switching frequency of 960 Hz. We note that these imaging parameters specifically optimize the signal in the inferior frontal region and may lead to worse signal in other areas such as the medial temporal regions (including the amygdala).

The matrix size was 64 × 64 and the field of view was

192 × 192 mm. Continuous coverage was obtained from +60 (A/P) to −87 (A/P). Acquisition was carried out during the task performance, which lasted a total of 30 min, yielding 600 volumes in total (not including the first 4 volumes that were discarded to allow for T1 equilibration). A whole-brain T2*-weighted EPI volume of the above dimensions and an anatomical T1 volume with slice thickness 7 mm and in-plane resolution of 0.75 × 0.75 mm was also acquired.

Data analysis

Image preprocessing was performed with FLIRT (FMRIB Linear Registration Tool, http://www.fmrib.ox.ac.uk/fsl/flirt) for realignment, reslicing with sinc interpolation, and normalisation to MNI coordinate system (Montreal Neurological Institute, Collins et al., 1994) used throughout this paper. SPM99 (http://www.fil.ion.ucl.ac.uk/spm) was used first to apply spatial smoothing with a 10-mm full width at half-maximum isotropic Gaussian kernel and global scaling. Time series at each voxel were then high-pass filtered and low-pass filtered with a haemodynamic response kernel.

We performed a conjunction analysis across the nine subjects to look for activations common to all subjects, where a conjunction was considered significant when the comparison was significant in every subject-specific model (Friston et al., 1999). Thus the P value corresponds to the null hypothesis of no activation in any subject. All nine subjects participated in the main reversal task and the face expression control task, while a subset of four subjects participated in the control reversal task which was subsequently used with the main reversal task in a conjunction analysis over eight datasets. The other subjects were unable to participate in the control reversal task due to time constraints. A threshold of $P < 0.05$ corrected for the entire brain was used for all comparisons and with this threshold we can infer that >70% of the population would have the same effect (Friston et al., 1999).

Modelling of conditions based on event-related responses

In the main reversal task and in the control reversal task, statistical parametric maps were generated in SPM99 based on each subject's event-related responses for each of two main types of events: acquisition (acq), and reversal (rev). *Acquisition* included series of trials where the subject's correct choices of a particular person led to a smile expression on that face (see Fig. 1, top). *Reversal* included individual trials where the subject had selected the previously rewarded person, but this had led to an angry face expression (see Fig. 1, bottom) in the main reversal task and a neutral face expression in the control reversal task. The trial that followed a reversal trial was modelled as an effect of no interest, so that the main statistical comparison would be between trials with stable acquisition of a smiling face, and reversal trials on which a change of face expression signalled reversal. The very occasional mistake (on average 3.2 of 85.8 trials per subject) was modelled as an event type of no interest. Each of the events was defined to include the trials irrespective of whether they were obtained from the stimulus set of female or male persons, and thus controlled for possible gender and face identification differences.

In the face expression control task two types of event were generated to create statistical parametric maps in SPM99 based on each subject's event-related responses: smile and angry face expression. The change to a smiling face expression in this control task was comparable to a smiling face expression obtained for the acquisition events in the main reversal task, while the angry face expression event in this control task was comparable to the angry face expression obtained in the main reversal task, but importantly without the contextual meaning that a change of behaviour was necessary.

Task-specific contrasts

In the main reversal task we were specifically interested in two contrasts: the effects of reversal [rev − acq] and the main effects of faces [rev AND acq]. We were particularly interested in the difference between conditions in which a face expression signals reversal and in which it does not. For each of the nine subjects we therefore established these two contrasts for the main reversal task, which were then used across all subjects in a conjunction analysis.

In the control reversal task we were interested in finding the common reversal component between this control reversal task and the main reversal task and, therefore, constructed a model including both these tasks in the four subjects who participated in both. We estimated the effects of reversal, and created a conjunction across the two tasks for this contrast. This conjunction thus revealed the areas commonly activated in reversal irrespective of which face expression was being used to cue the reversal.

We also established the effects of smiling and angry face expression in the face expression control task for each subject. The results of these analyses were also used in a conjunction analysis across all subjects to reveal the main effects of face expression.

Time courses of activations

To further establish the effects of changes in face expression on three different brain areas—the fusiform gyrus, the orbitofrontal cortex, and the anterior cingulate/paracingulate cortex—we used the time courses of the activations from the two main contrasts to establish the elicited time course for each type of event in each of these brain areas in each of the three tasks. In other words, for each type of event we measured the mean change in blood oxygenation level-dependent (BOLD) changes over time in a cluster of voxels. The time courses of activation allowed us to make inferences about the significance of changes in face expression for the activity in different brain regions. Specifically we were able to compare the effects of angry face expression on brain activity in a reversal and nonreversal context

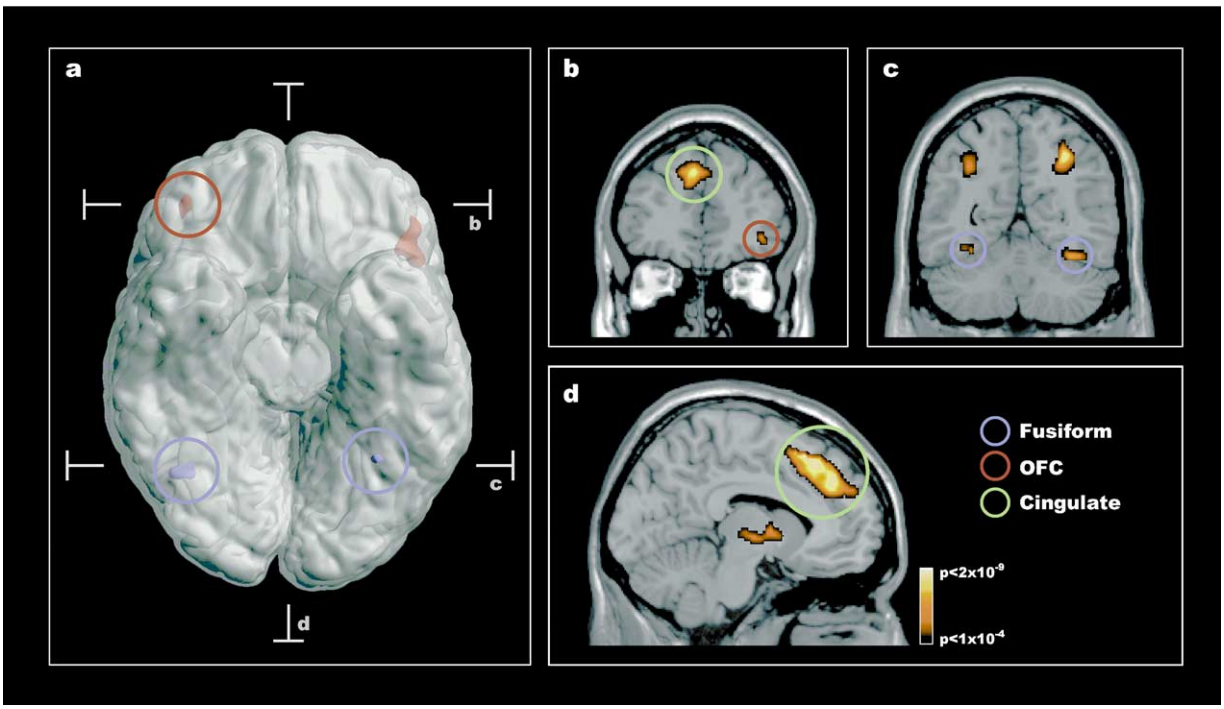


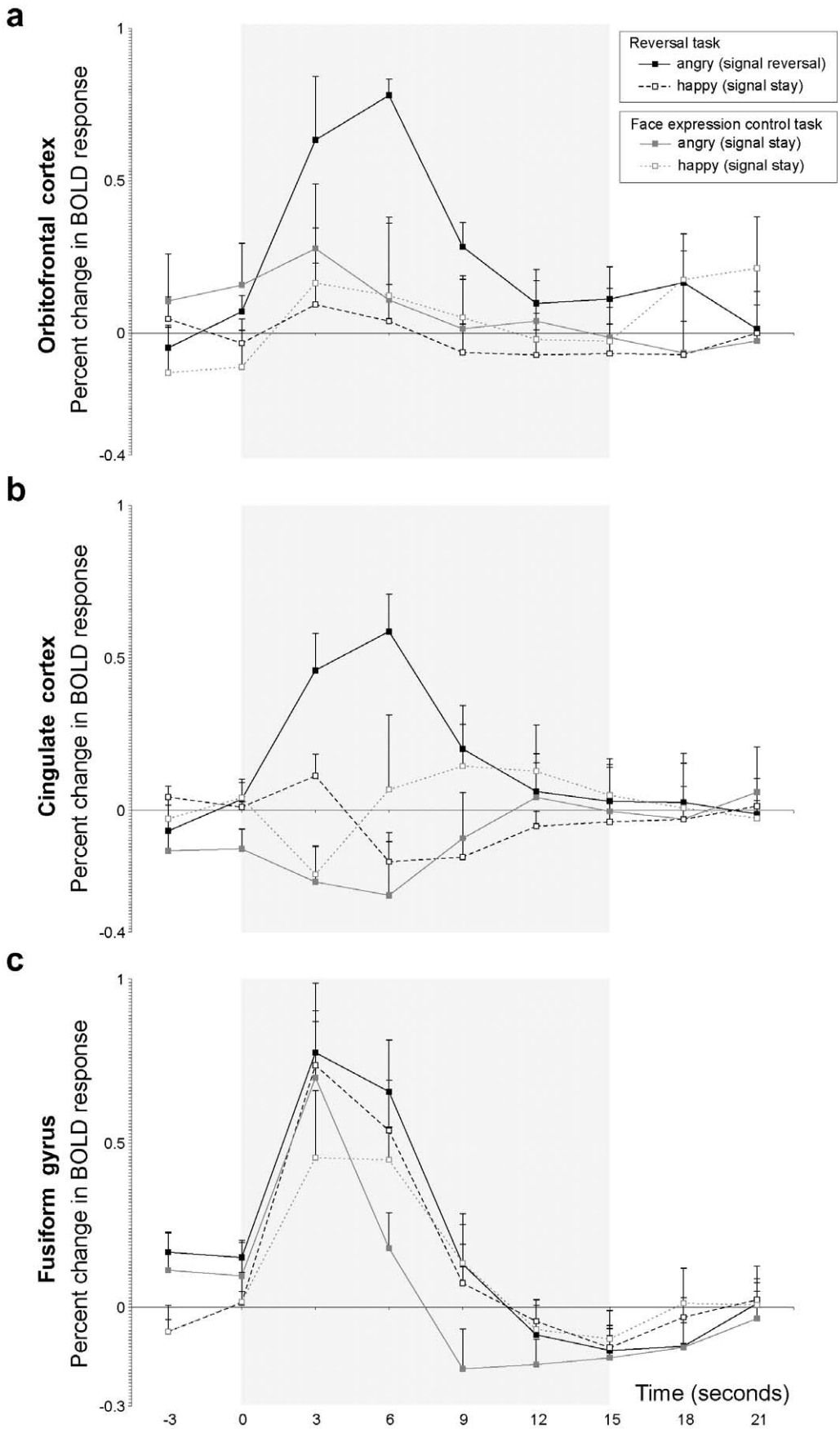
Fig. 2. Composite figure showing that changing behaviour based on face expression is correlated with increased brain activity in the human orbitofrontal cortex. (a) The figure is based on two different group statistical contrasts from the neuroimaging data which are superimposed on a ventral view of the human brain with the cerebellum removed, and with indication of the location of the two coronal slices (b,c) and the transverse slice (d). The red activations in the orbitofrontal cortex (denoted *OFC*, maximal activation: $Z = 4.94$; 42, 42, −8, and $Z = 5.51$; *x*, *y*, *z* = 46, 30, −8) shown on the rendered brain arise from a comparison of reversal events with stable acquisition events, while the blue activations in the fusiform gyrus (denoted *Fusiform*, maximal activation: $Z > 8$; 36, −60, −20, and $Z = 7.80$; −30, −56, −16) arise from the main effects of face expression. (b) The coronal slice through the frontal part of the brain shows the cluster in the right orbitofrontal cortex across all nine subjects when comparing reversal events with stable acquisition events. Significant activity was also seen in an extended area of the anterior cingulate/paracingulate cortex (denoted *Cingulate*, maximal activation: $Z = 6.88$; −8, 22, 52; green circle). (c) The coronal slice through the posterior part of the brain shows the brain response to the main effects of face expression with significant activation in the fusiform gyrus and the cortex in the intraparietal sulcus (maximal activation: $Z > 8$; 32, −60, 46, and $Z > 8$: −32, −60, 44). (d) The transverse slice shows the extent of the activation in the anterior cingulate/paracingulate cortex when comparing reversal events with stable acquisition events. Group statistical results are superimposed on a ventral view of the human brain with the cerebellum removed, and on coronal and transverse slices of the same template brain (activations are thresholded at $P = 0.0001$ for purposes of illustration to show their extent).

(by comparing the time course of activations to angry face expression in the main reversal task with those in the face expression control task, see Fig. 3). Furthermore we were able to compare the effects of different face expressions signalling reversal by comparing the reversal cue in the main reversal task (an angry face expression) with the reversal cue in the control reversal task (a neutral face expression, see Fig. 4).

Results

To analyse the neural correlates of reversal learning we divided the main reversal task into two main types of event that depended on the subjects' performance. The first kind of event, acquisition, included series of trials where the subject's correct choices of a particular person led to a smile expression on that face (see Fig. 1, top, and Materials and Methods). The second kind of event, reversal, included individual trials where the subject had selected the previously rewarded person, but this had led to an angry face expression (see Fig. 1, bottom). This reversal type of trial led the previously trained subjects to change their behavioural choice and to select the other person's face on the next trial, which then smiled.

The principal comparison in the main reversal task was made between the trials where the subject selected the "happy" person (and therefore received a smile) and those where the subject received an angry face expression signalling a change in mood, and thus that a change in behaviour was required. The group activations for the reversal minus acquisition contrast showed significant effects in the orbitofrontal cortex (maximal activation: $Z = 5.51$; *x*, *y*, $z = -46$, 30, −8, and $Z = 4.94$: 42, 42, −8, red circles in Figs. 2a and b and Figs. 3a), and in the anterior cingulate/paracingulate cortex (maximal activation: $Z = 6.88$; *x*, *y*, $z = -8$, 22, 52; see Figs. 2b and d and Figs. 3b and 4b) with the extent of the activation covering an area from $y = 12$ to $y = 54$ and from $z = 55$ to $z = 25$ (see Fig. 2d). Activation was also found in the bilateral inferior precentral sulcus (BA44/45) (maxi-


a
b
c
Orbitofrontal cortex
Cingulate cortex
Fusiform gyrus
Percent change in BOLD response
1
0.5
0
-0.4
-0.3
Reversal task
angry (signal reversal)
happy (signal stay)
Face expression control task
angry (signal stay)
happy (signal stay)
Time (seconds)
-3
0
3
6
9
12
15
18
21

mal activation: $Z = 5.31$; $x, y, z = 38, 20, 16$; maximal activation: $Z = 4.80$; $x, y, z = -46, 16, 16$). Furthermore activation was found in the region of ventral striatum ($Z = 5.08$; $x, y, z = -10, -10, 0$; $Z = 5.09$; $x, y, z = -16, 8, 4$). The activations are summarized in Table 1.

Interestingly, this contrast did not show significant effects in the posterior brain areas, including those activated by faces which included the fusiform gyrus (maximal activation to faces: $Z > 8$; $x, y, z = 36, -60, -20$, and $Z = 7.80$; $x, y, z = -30, -56, -16$; blue circles in Fig. 2a) and the cortex in the intraparietal sulcus (maximal activation to faces: $Z > 8$; $x, y, z = 32, -60, 46$, and $Z > 8$; $x, y, z = -32, -60, 44$; Fig. 2c). This was because the elicited activations were comparable in magnitude for both reversal and acquisition events (see below and compare the time courses of activation to the different events in Fig. 3c). These posterior areas were in other words robustly activated by faces, and the activations in the fusiform gyrus were very similar both for angry faces when they signalled a reversal and for happy faces (see Fig. 3c).

To investigate whether the response to the angry face was linked to the angriness of the face expression, we used the face expression control task (see Materials and methods). This control task required the subjects to look at the screen and to make choices, but there was no reversal contingency component to this task. The main effects of angry face expression in the face expression control task (where no reversal was required to this face expression) showed robust activation of the fusiform gyrus (and the intraparietal cortex) to the angry face expression, but no significant activation in the orbitofrontal and cingulate cortical areas activated during the reversal.

To further investigate the effects of angry face expression in the orbitofrontal and in the anterior cingulate/paracingulate cortices we used the clusters obtained in the reversal contrast in the reversal task, and extracted the time courses of activation for each event in the reversal task and in the face expression control task. As is shown by the time courses in Figs. 3a and b, the responses in the orbitofrontal and cingulate/paracingulate cortices were large to the angry face only in the reversal task in which the angry face signalled that behaviour should change, and not in the face expression control task. Moreover, the activations in the orbitofrontal and cingulate cortices to the happy face in the reversal and face expression control tasks were similar (Figs. 3a and b), providing evidence that these cortical areas did not simply become generally more responsive to faces in the experimental task, but in fact showed strong activation only when an angry face indicated that the behaviour should change.

To obtain even further evidence on the linking of the activations in the orbitofrontal and cingulate/paracingulate cortices to the change of behaviour that was required in the reversal task (rather than just to angry face expression per se), we used the control reversal task (see Materials and methods) and ran this task in a subset of the nine subjects. In this control reversal task reversal was cued by a change to a neutral rather than to an angry face expression, and given that the structures of both the main reversal task and the control reversal were identical with only the reinforcing face expressions being different, we were able to compare them directly in each of the subjects who performed both tasks (see Fig. 4).

The reversal–acquisition contrast in the control reversal task again showed significant activations in the orbitofrontal and anterior cingulate/paracingulate cortices overlapping with those found in the main reversal task. We were then able to directly compare the responses to reversal learning *per se* by creating a model that included both the main reversal task and the control reversal task in the subjects that performed both tasks (see Materials and methods). We could therefore create a conjunction contrast to investigate the overlap between the reversal–acquisition events in each of the two tasks. Again we found activation in very similar areas of the orbitofrontal cortex bilaterally (maximal activation: $Z = 5.39$: $x, y, z = 44, 26, -12$, and $Z = 4.33$; $x, y, z = -46, 28, -8$) and a large area of the anterior cingulate/paracingulate cortex (maximal activation: $Z = 7.55$; $x, y, z = -6, 18, 48$). In addition we also found activations in areas of the bilateral inferior precentral sulcus (BA44/45)

Fig. 3. Group time courses for (a) the orbitofrontal cortex, (b) the anterior cingulate/paracingulate cortex, and (c) the fusiform gyrus, across all subjects. To uncover how different brain areas were modulated by different face expressions depending on the context, we plotted the blood oxygenation level-dependent changes in different brain areas over time in the reversal and face expression control tasks. In the reversal task the angry face expression (in black) was always associated with a behavioural change, while a happy face (in stippled black) indicated stable acquisition. This was in contrast to the face expression control task where an angry face expression (in grey) did not signal a behavioral change, as the subjects were instructed to find the person whose face expression changed randomly (happy face expressions in stippled grey), and to keep choosing this person, while the other person remained unresponsive. (a) It is only when the angry face expression indicates that a behavioural change is needed (and not just when an angry face expression is shown) that the human orbitofrontal cortex is strongly recruited (compare time courses of the black line from the reversal task with the other three lines). The orbitofrontal cortex is significantly more activated after 6 s by angry face expressions signalling reversal than by angry face expressions not signalling reversal, while happy faces do not appear to activate this part of the orbitofrontal cortex. The time courses associated with each type of face expression were averaged over all subjects from the significantly activated cluster of five voxels in the orbitofrontal cortex ($-46, 30, -8$) from the [reversal − acquisition] contrast. (b) A very similar effect was seen in the anterior cingulate/paracingulate cortex which was also significantly activated by angry face expressions signalling reversal and not by angry face expressions not signalling reversal. The time courses associated with each type of face expression were averaged over all subjects from the significantly activated cluster of eight voxels in the anterior cingulate cortex ($-8, 22, 52$) from the [reversal − acquisition] contrast. (c) The fusiform gyrus was activated bilaterally by all face expressions regardless of whether the face expression was angry or happy (compare the time courses of the four lines). The time courses associated with each type of face expression were averaged over all subjects from the significantly activated cluster of 15 voxels in the fusiform gyrus ($-30, -56, -16$) in the contrast revealing the main effects of faces. The bars indicate the standard errors of the means across subjects.

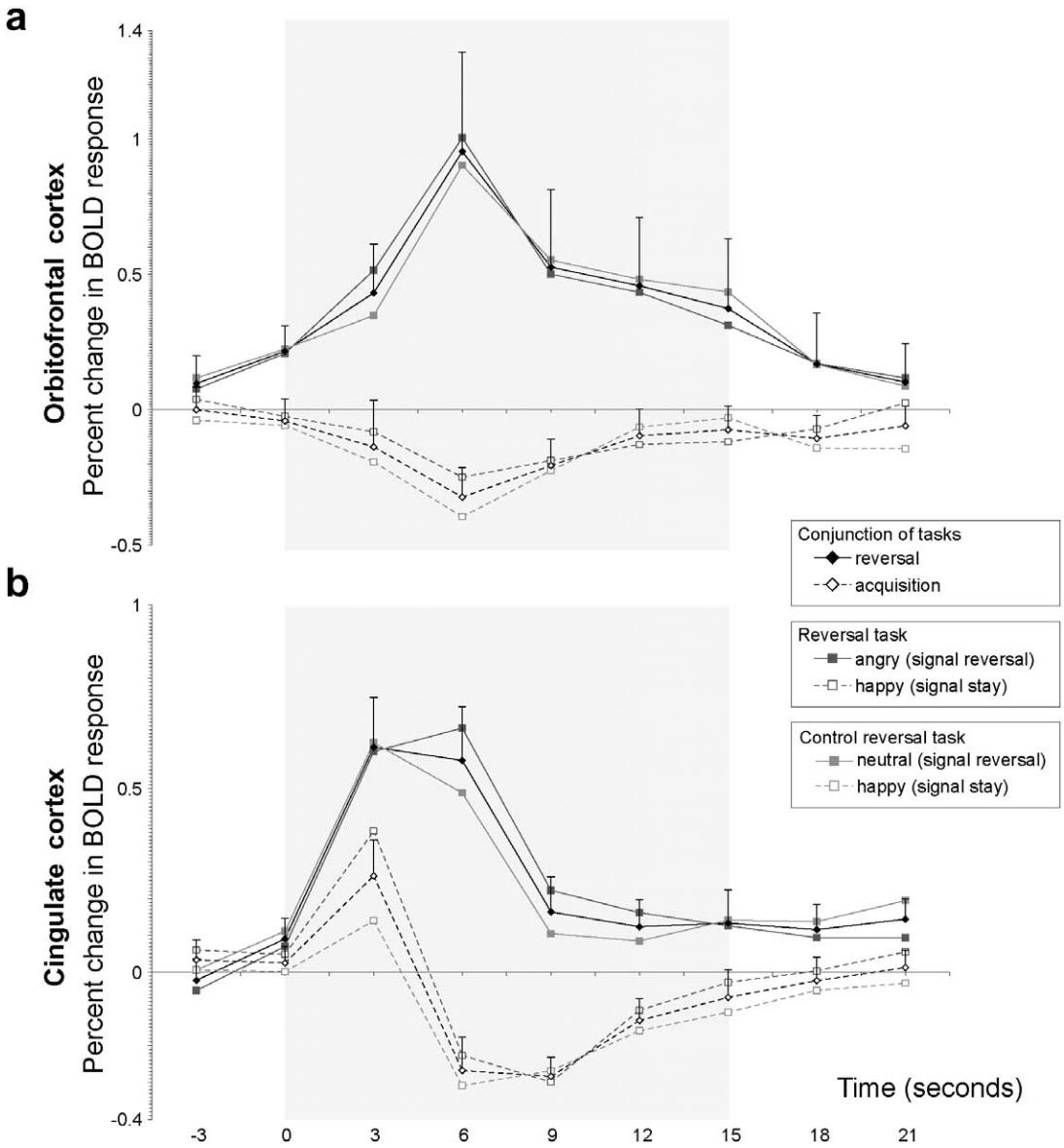


Fig. 4. Group time courses of activations are shown for (a) the orbitofrontal cortex and (b) the anterior cingulate/paracingulate cortex across subjects having performed both the main reversal task and the control reversal task. We created a conjunction of the two reversal tasks to uncover what brain regions were activated by the reversal irrespective of the face expressions used. Two brain areas, the orbitofrontal and anterior cingulate/paracingulate cortices, were consistently more activated during the reversal events (black) than during stable acquisition (black stippled), regardless of whether an angry or a neutral face expression signalled a change in behaviour (in the reversal and control reversal tasks, respectively). (a) The figure shows the time courses of the activations produced in the main reversal task by angry faces (dark grey) and happy faces (dark grey stippled), and in the control reversal by neutral faces (light grey) and happy faces (light grey stippled) in the cluster of 18 voxels in the orbitofrontal cortex (maximal activation: $Z = 5.39$; −44, 26, −12) overlapping with the cluster described in Fig. 3a. There was an additional cluster in the orbitofrontal cortex (maximal activation: $Z = 4.33$; 46, 28, −8). (b) This effect was mirrored by the time courses of a cluster of 17 voxels in the anterior cingulate cortex (maximal activation: $Z = 7.55$; −6, 18, 48) in a cluster very similar to that described in Fig. 3b. The bars indicate the standard errors of the means across subjects.

(maximal activation: $Z = 6.78$: $x, y, z = 40, 18, 28$, and $Z = 5.95$; $x, y, z = -46, 10, 36$).

In addition, we were again able to compare the time courses from the control reversal task with the time courses from the main reversal task in the same subjects. As shown in Fig. 4 the results clearly demonstrate that irrespective of the reinforcing stimuli (i.e., whether using angry or neutral face expression as a cue to reversal in the reversal task and the control reversal task, respectively), both the orbitofrontal and cingulate/paracingulate cortices are specifically activated when reversing behaviour in this control reversal task, as well as in the main reversal task, compared with the happy face expression indicating stable acquisition. This result strongly indicates that the activations in both the orbitofrontal and cingulate/paracingulate cortices are related to the reversal cued by a change of face expression and not

Table 1
Brain regions involved in reversal learning[a]

		MNI *x*	MNI *y*	MNI *z*	*Z* score
Main reversal task					
Reversal [rev − acq]					
Orbitofrontal cortex	Right	42	42	−8	4.94
	Left	−46	30	−8	5.51
Inferior precentral sulcus (BA44/45)	Right	38	20	16	5.31
	Left	−46	16	16	4.80
Ventral striatum	Left	−10	−10	0	5.08
	Left	−16	8	4	5.09
Anterior cingulate/paracingulate cortex	Left	−8	22	52	6.88
Main effects of faces [rev AND acq]					
Fusiform cortex	Right	36	−60	−20	>8
	Left	−30	−56	−16	7.8
Intraparietal sulcus	Right	32	−60	46	>8
	Left	−32	−60	44	>8
Conjunction of main reversal task and control reversal task					
Reversal [rev − acq]					
Orbitofrontal cortex	Right	44	26	−12	5.39
	Left	−46	28	−8	4.33
Anterior cingulate/paracingulate cortex	Left	−6	18	48	7.55
Inferior precentral sulcus (BA44/45)	Right	40	18	28	6.78
	Left	−46	10	36	5.95

[a] All activations are significant at $P < 0.05$, corrected for the entire brain volume.

to any particular face expression. This links these brain regions to a general mechanism involved in reversal learning of which the reversal involved in human social interaction is an important example.

Discussion

Our results provide the first neuroimaging evidence that the human orbitofrontal and anterior cingulate/paracingulate cortices are part of a general mechanism for reversal learning and, furthermore, indicate that these brain areas are also important in changing behaviour in a simple model of social interactions based on changes in face expression. Human social behaviour is obviously very complex, and reversal learning is a small but arguably very important part of normal human social interaction. The evidence presented here demonstrates that the activity in these brain regions is produced by the face expression signalling behavioural change rather than by any particular face expression. Moreover, as shown in Figs. 3a, 3b, 4a, and 4b, these regions of the orbitofrontal and anterior cingulate cortices are active only when a face expression is signalling reversal, and not when the same face expression is shown in another context in which reversal is not required. The results show that the same regions of the orbitofrontal and anterior cingulate cortices are activated when reversing regardless of whether angry and neutral face expressions are provided as the cue to reversing behaviour, and this indicates that these brain regions are involved in a general mechanism for reversal learning that is not restricted to particular face expressions.

The design of the current study enabled us for the first time to explicitly image in relation to the actual event that produced reversal. This contrasts with a probabilistic design for the magnitudes of the monetary rewards and punishers used in a previous study (O'Doherty et al., 2001), which enabled us to analyse brain activations related to magnitudes of reinforcers, but not to the actual change in behaviour itself. The current study shows directly the brain regions related to the reversal events. Moreover, the current study shows that regions of the orbitofrontal cortex and a region located on the border between the anterior cingulate and paracingulate cortex are activated in relation to the reversal event. Given that there are direct projections from the orbitofrontal cortex downstream to the anterior cingulate/paracingulate cortex in non-human primates (Öngür and Price, 2000), we suggest that during reversal, these two structures operate as a linked pair. In particular, there is much evidence that the orbitofrontal cortex performs the function of evaluating the reward and punishment value of reinforcers (Rolls, 1999).

The significance of the co-activation in the region on the border between the anterior cingulate and paracingulate cortex with the orbitofrontal cortex opens a number of interesting interpretations depending on whether this group activation is in the anterior cingulate cortex proper or anterior paracingulate cortex/medial superior frontal gyrus. However, the correct cortical localisation of group activation in neuroimaging studies is notoriously difficult to establish on the medial wall of the human brain where there is high variability from person to person in the pattern of cortical folding (Paus et al., 1996b). The paracingulate sul-

cus is present in only 30–50% of individuals, and the volume (Paus et al., 1996a) and the incidence (Ide et al., 1999; Paus et al., 1996b; Yucel et al., 2001) of the paracingulate sulcus are greater in the left hemisphere.

If the group activation observed in this study is indeed in the anterior paracingulate cortex, this would be further evidence for a role of this region in theory of mind and mentalizing (Adolphs, 2003; Blakemore and Decety, 2001; Frith and Frith, 1999; Gallagher and Frith, 2003). Previous neuroimaging studies have implicated the paracingulate cortex in a task that involved attributing other beliefs and goals to other people compared with a machine or random sequence (Gallagher et al., 2002), and also in a task that involved inferring theory of mind in verbal and nonverbal tasks (Gallagher et al., 2000). The peak of activations found in these and other studies (see review Gallagher and Frith, 2003) are within the cluster of activation in the anterior paracingulate cortex found in the present study. Also of potential interest to this line of interpretation is the important evolutionary evidence that the cingulate region has an unusual type of projection neuron (spindle cell) found only in humans and in some other higher primates (hominids and pongids), but not in monkeys (Nimchinsky et al., 1999). Interestingly, in humans these cells are not present at birth, but appear at around 4 months of age (Allman et al., 2001).

Alternatively, if the group activation related to reversal is in the anterior cingulate cortex proper, the signal from the orbitofrontal cortex may enable the anterior cingulate cortex to perform a function in response selection, consistent with other neuroimaging studies that have shown that this region is activated in, for example, Stroop tasks (Carter et al., 1995), reward-based decision-making tasks (Bush et al., 2002), anticipatory arousal (Critchley et al., 2001), or cognitive demand (Duncan and Owen, 2000). Another possibility is that the anterior cingulate cortex serves an evaluative function in conflict monitoring as suggested by other neuroimaging studies (Botvinick et al., 1999; Carter et al., 1998). This would be compatible with a role for the anterior cingulate cortex in the evaluation of the conflict between an established stimulus–response mapping and the updating of this at the reversal trial, and might suggest that the anterior cingulate cortex could be part of an "error prevention" network (Carter et al., 2000).

The findings reported here are consistent with the essential role of the human orbitofrontal cortex in reversal tasks as demonstrated by orbitofrontal cortex lesions in nonhuman primates (Dias et al., 1996; Iversen and Mishkin, 1970; Meunier et al., 1997) and in humans with damage to the ventral surface of the brain (Rolls et al., 1994) including the orbitofrontal cortex (Hornak et al., 2003b). It is also consistent with the evidence from the effects of lesions in humans that the orbitofrontal cortex is important for social response reversal (Blair and Cipolotti, 2000). Furthermore, single-neuron recordings in macaques have found activity in the orbitofrontal cortex when making associations between a previous neutral (e.g., visual) stimulus and a (typically primary) reinforcer, and then rapidly reversing these associations when the reinforcement contingencies alter (Thorpe et al., 1983). Thorpe et al. (1983) showed that not only do some orbitofrontal cortex neurons reverse the visual stimulus to which they respond when the reward association reverses, but also, in a direct parallel to what is reported in the present fMRI study, that a separate population of orbitofrontal cortex neurons respond to a mismatch between the expected reward and the reinforcer actually obtained. These latter neurons are thus error detection neurons. The findings described here also raise the interesting possibility that similar changes in social behaviour could be produced by damage to either the human orbitofrontal or anterior cingulate/paracingulate cortex, because they are part of a linked system that is coactive when behaviour and in particular social behaviour must change. Indeed, changes in social and emotional behaviour have now been demonstrated in patients with discrete surgical lesions of either the orbitofrontal cortex or the anterior cingulate cortex (Hornak et al., 2003b).

The reversal-related activity in the inferior precentral sulcus both in the main reversal task and in the control reversal task is consistent with similar results in a number of other recent imaging studies featuring tasks associated with prefrontal function, including a probabilistic monetary gambling task (O'Doherty et al., 2001), the cognitive set-shifting component of the Wisconsin card sorting task (Konishi et al., 1998a), and the no-go trials of a go/no-go task (Konishi et al., 1998b). These findings suggest when taken together with the present finding that the cortex in the inferior prefrontal sulcus may be involved in inhibiting inappropriate behavioural strategies.

The evidence that the ventral striatum is significantly activated in the main reversal task is consistent with the known direct anatomical connections from the orbitofrontal cortex in a segregated frontostriatal loop (Alexander et al., 1986). The ventral striatum has been proposed to relay motivational information from the limbic to the motor system, which mediates the effects of stimulus–reward mechanisms on goal-directed behaviour (Robbins et al., 1989; Schultz et al., 1992). Recent human imaging studies have implicated the ventral striatum in many reward-related tasks, e.g., in probabilistic reversal learning (Cools et al., 2002) and in the expectation to pleasant taste (O'Doherty et al., 2002). It has also recently been shown that administration of dopaminergic medication to patients with mild Parkinson's disease has a detrimental effect on probabilistic reversal learning by effectively overdosing a relatively intact ventral frontostriatal system (Cools et al., 2001; Swainson et al., 2000).

In the reversal and control tasks the activations in the fusiform gyrus to both happy and angry face expressions were independent of whether a reversal was to be performed. This leads to the interesting conclusion that the fusiform face area represents face expression independently of whether it signals that behaviour should change, whereas

the orbitofrontal and cingulate cortices are activated specifically when the face expression indicates that a behavioural change is needed. The similar activations of the fusiform face area in the reversal and control tasks are consistent with other studies implicating these areas in face perception and attention (Haxby et al., 1994; Kanwisher, 2000), and provide evidence that the subjects were attending to the faces, which was also reflected by their one-trial reversal behaviour in the reversal task. The stringent statistical threshold used in this investigation (conjunction analysis at $P < 0.05$ corrected for multiple comparisons) did not reveal activity in the posterior superior temporal sulcus, which other neuroimaging studies have found is sensitive to biological motion (Grossman and Blake, 2001) and facial movements (Puce et al., 1998; Rizzolatti et al., 1996; Wicker et al., 1998). However, if we lower the statistical threshold when comparing reversal with acquisition events, we do observe activity in the posterior superior temporal sulcus ($Z = 3.43$; $x, y, z = -58, -50, 26$, $P < 0.001$, uncorrected). This lower significance could be due to our special imaging parameters which optimise the signal in the orbitofrontal cortex but lead to weaker signal in other brain areas. A case in point is the amygdala where other neuroimaging studies have found activation to face expression, and the absence of amygdala activation in this study is likely to be linked to our special imaging parameters which leads to weaker signal in regions such as the anterior medial temporal regions including the amygdala.

Previous neuroimaging investigations have implicated a number of brain structures in the decoding of face expression as distinct from face perception and identification, but none of these previous studies has investigated the effects of using changes in face expression in a context-dependent naturalistic learning situation as a cue to alter behaviour. The results of this study thus provide the first neuroimaging evidence that the human orbitofrontal and anterior cingulate/paracingulate cortices are important for changing behaviour in social interactions. This is consistent with the role of the human orbitofrontal cortex in processing both primary rewards (e.g., Kringelbach et al., 2003) and more abstract rewards such as money (O'Doherty et al., 2001; Thut et al., 1997). Indeed, the orbitofrontal cortex has been proposed to be important for emotional processing in general (Damasio, 1994; Rolls, 1990, 1999), and this study provides important new evidence on how this brain region is a driving force in shaping behaviour based on subtle social cues, and on how this brain region together with the anterior cingulate/paracingulate cortex is an important part of a general mechanism for reward reversal learning.

Acknowledgments

This research is supported by the Medical Research Council (to FMRIB, to the IRC for Cognitive Neuroscience, and by Programme Grant 9826105 to E.T.R.); the Danish Research Agency, Aarhus, Denmark (to M.L.K.); the Oxford McDonnell–Pew Centre for Cognitive Neuroscience; and by Firmenich S.A. We thank Ivan de Araujo, Peter Hobden, Christopher Wright, and Håkan Fischer for their help and assistance.

References

Adolphs, R., 2003. Cognitive neuroscience of human social behaviour. Nat. Rev. Neurosci. 4, 165–178.

Adolphs, R., Tranel, D., Damasio, H., Damasio, A., 1994. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature 372, 669–672.

Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu. Rev. Neurosci. 9, 357–381.

Allman, J.M., Hakeem, A., Erwin, J.M., Nimchinsky, E., Hof, P., 2001. The anterior cingulate cortex: the evolution of an interface between emotion and cognition. Ann. NY Acad. Sci. 935, 107–117.

Blair, R.J., Cipolotti, L., 2000. Impaired social response reversal: a case of 'acquired sociopathy'. Brain 123, 1122–1141.

Blakemore, S.J., Decety, J., 2001. From the perception of action to the understanding of intention. Nat. Rev. Neurosci. 2, 561–567.

Bodamer, J., 1947. Die Prosop-Agnosie. Arch. Psychiatrie Nervenkrankheiten 179, 6–53.

Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402, 179–181.

Bruce, C., Desimone, R., Gross, C.G., 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J. Neurophysiol. 46, 369–384.

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A., et al., 2002. Dorsal anterior cingulate cortex: a role in reward-based decision making. Proc. Natl. Acad. Sci. USA 99, 523–528.

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D., 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280, 747–749.

Carter, C.S., Macdonald, A.M., Botvinick, M., Ross, L.L., Stenger, V.A., Noll, D., et al., 2000. Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. Proc. Natl. Acad. Sci. USA 97, 1944–1948.

Carter, C.S., Mintun, M., Cohen, J.D., 1995. Interference and facilitation effects during selective attention: an $H_2{}^{15}O$ PET study of Stroop task performance. NeuroImage 2, 264–272.

Collins, D., Neelin, P., Peters, T., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assisted Tomogr. 18, 192–205.

Cools, R., Barker, R.A., Sahakian, B.J., Robbins, T.W., 2001. Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. Cereb. Cortex 11, 1136–1143.

Cools, R., Clark, L., Owen, A.M., Robbins, T.W., 2002. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. J. Neurosci. 22, 4563–4567.

Critchley, H.D., Mathias, C.J., Dolan, R.J., 2001. Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron 29, 537–545.

Damasio, A.R., 1994. Descartes' error, Putnam, New York.

Darwin, C., 1872. The Expression of the Emotions in Man and Animals. Univ. of Chicago Press, Chicago.

Davidson, R.J., Putnam, K.M., Larson, C.L., 2000. Dysfunction in the neural circuitry of emotion regulation a possible prelude to violence. Science 289, 591–594.

Desimone, R., Gross, C.G., 1979. Visual areas in the temporal cortex of the macaque. Brain Res. 178, 363–380.

Dias, R., Robbins, T., Roberts, A., 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 Neurosci. 23, 475–483.

Ekman, P., Friesen, W-V., 1971. Constants across cultures in the face and emotion. J. Pers. Social Psychol. 17, 124–129.

Eslinger, P., Damasio, A., 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. Neurology 35, 1731–1741.

Friston, K.J., Holmes, A.P., Price, C.J., Buchel, C., Worsley, K.J., 1999. Multisubject fMRI studies and conjunction analyses. NeuroImage 10, 385–396.

Frith, C.D., Frith, U., 1999. Interacting minds: a biological basis. Science 286, 1692–1695.

Gallagher, H.L., Frith, C.D., 2003. Functional imaging of 'theory of mind'. Trends Cogn. Sci. 7, 77–83.

Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia 38, 11–21.

Gallagher, H.L., Jack, A.I., Roepstorff, A., Frith, C.D., 2002. Imaging the intentional stance in a competitive game. NeuroImage 16, 814–821.

Grossman, E.D., Blake, R., 2001. Brain activity evoked by inverted and imagined biological motion. Vision Res. 41, 1475–1482.

Hasselmo, M.E., Rolls, E.T., Baylis, G.C., 1989a. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behav. Brain Res. 32, 203–21.

Hasselmo, M.E., Rolls, E.T., Baylis, G.C., Nalwa, V., 1989b. Object-centred encoding by face-selective neurons in the cortex in the superior temporal sulcus of the the monkey. Exp. Brain Res. 75, 417–429.

Haxby, J.V., Grady, C.L., Horwitz, B., Ungerleider, L.G., Mishkin, M., Carson, R.E., et al., 1991. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. Proc. Natl. Acad. Sci. USA 88, 1621–1625.

Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., Grady, C.L., 1994. The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J. Neurosci. 14, 6336–6353.

Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R., et al., 2003a. Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. Brain 126, 1671–1712.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., 2003b. Reward-related reversal learning after surgical excisions in orbitofrontal and dorsolateral prefrontal cortex in humans. J. Cogn. Neurosci. in press..

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.

Ide, A., Dolezal, C., Fernandez, M., Labbe, E., Mandujano, R., Montes, S., et al., 1999. Hemispheric differences in variability of fissural patterns in parasylvian and cingulate regions of human brains. J. Comp. Neurol. 410, 235–242.

Iversen, S.D., Mishkin, M., 1970. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp. Brain Res. 11, 376–386.

Kanwisher, N., 2000. Domain specificity in face perception. Nat. Neurosci. 3, 759–763.

Köhler, S., Kapur, S., Moscovitch, M., Winocur, G., Houle, S., 1995. Dissociation of pathways for object and spatial vision: a PET study in humans. NeuroReport 6, 1865–1868.

Konishi, S., Nakajima, K., Uchida, I., Kameyama, M., Nakahara, K., Sekihara, K., et al., 1998a. Transient activation of inferior prefrontal cortex during cognitive set shifting. Nat. Neurosci. 1, 80–84.

Konishi, S., Nakajima, K., Uchida, I., Sekihara, K., Miyashita, Y., 1998b. No-go dominant brain activity in human inferior prefrontal cortex revealed by functional magnetic resonance imaging. Eur. J. Neurosci. 10, 1209–1213.

Kringelbach, M.L., O'Doherty, J., Rolls, E.T., Andrews, C., 2003. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. Cereb. Cortex, in press.

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.

Nimchinsky, E.A., Gilissen, E., Allman, J.M., Perl, D.P., Erwin, J.M., Hof, P.R., 1999. A neuronal morphologic type unique to humans and great apes. Proc. Natl. Acad. Sci. USA 96, 5268–5273.

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

O'Doherty, J.P., Deichmann, R., Critchley, H.D., Dolan, R.J., 2002. Neural responses during anticipation of a primary taste reward. Neuron 33, 815–826.

Öngür, D., Price, J.L., 2000. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb. Cortex 10, 206–219.

Paus, T., Otaky, N., Caramanos, Z., MacDonald, D., Zijdenbos, A., D'Avirro, D., et al., 1996a. In vivo morphometry of the intrasulcal gray matter in the human cingulate, paracingulate, and superior-rostral sulci: hemispheric asymmetries, gender differences and probability maps. J. Comp. Neurol. 376, 664–673.

Paus, T., Tomaiuolo, F., Otaky, N., MacDonald, D., Petrides, M., Atlas, J., et al., 1996b. Human cingulate and paracingulate sulci: pattern, variability, asymmetry, and probabilistic map. Cereb. Cortex 6, 207–214.

Perrett, D.I., Rolls, E.T., Caan, W., 1982. Visual neurons responsive to faces in the monkey temporal cortex. Exp. Brain Res. 47, 329–342.

Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. J. Neurosci. 18, 2188–2199.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al., 1996. Localization of grasp representations in humans by PET: 1. Observation versus execution. Exp. Brain Res. 111, 246–252.

Robbins, T.W., Cador, M., Taylor, J.R., Everitt, B.J., 1989. Limbic–striatal interactions in reward-related processes. Neurosci. Biobehav. Rev. 13, 155–162.

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

Rolls, E.T., 1999. The Brain and Emotion, Oxford Univ. Press, Oxford.

Rolls, E.T., Deco, G. 2002. Computational Neuroscience of Vision, Oxford Univ. Press, Oxford.

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. J. Neurol. Neurosurg. Psychiatry 57, 1518–1524.

Schultz, W., Apicella, P., Scarnati, E., Ljungberg, T., 1992. Neuronal activity in monkey ventral striatum related to the expectation of reward. J. Neurosci. 12, 4595–4610.

Sergent, J., Villemure, J.G., 1989. Prosopagnosia in a right hemispherectomized patient. Brain 112, 975–995.

Swainson, R., Rogers, R.D., Sahakian, B.J., Summers, B.A., Polkey, C.E., Robbins, T.W., 2000. Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. Neuropsychologia 38, 596–612.

Thorpe, S.J., Rolls, E.T., Maddison, S., 1983. Neuronal activity in the orbitofrontal cortex of the behaving monkey. Exp. Brain Res. 49, 93–115.

Thut, G., Schultz, W., Roelcke, U., Nienhusmeier, M., Missimer, J., Maguire, R.P., et al., 1997. Activation of the human brain by monetary reward. NeuroReport 8, 1225–1228.

deWaal, F.B.M., 1997. Good Natured: The Origins of Right and Wrong in Humans and Other Animals. Harvard Univ. Press, Cambridge, MA.

Whiten, A., Byrne, R.W., 1997. Machiavellian Intelligence, 2nd ed. Cambridge Univ. Press, Cambridge, London/New York.

Wicker, B., Michel, F., Henaff, M.A., Decety, J., 1998. Brain regions involved in the perception of gaze: a PET study. NeuroImage 8, 221–227.

Wilson, J., Jenkinson, M., de Araujo, I.E.T., Kringelbach, M.L., Rolls, E.T., Jezzard, P., 2002. Fast, fully automated global and local magnetic field optimization for fMRI of the human brain. NeuroImage 17, 967–976.

Yucel, M., Stuart, G.W., Maruff, P., Velakoulis, D., Crowe, S.F., Savage, G., et al., 2001. Hemispheric and gender-related differences in the gross morphology of the anterior cingulate/paracingulate cortex in normal volunteers: an MRI morphometric study. Cereb. Cortex 11, 17–25.