The orbitofrontal cortex

EDMUND T. ROLLS

University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, U.K.

SUMMARY

The orbitofrontal cortex contains the secondary taste cortex, in which the reward value of taste is represented. It also contains the secondary and tertiary olfactory cortical areas, in which information about the identity and also about the reward value of odours is represented. The orbitofrontal cortex also receives information about the sight of objects from the temporal lobe cortical visual areas, and is involved in learning and in reversing stimulus-reinforcement associations. The stimulus might be a visual or olfactory stimulus, and the primary (unlearned) reinforcer a taste or touch. Damage to the orbitofrontal cortex impairs the learning and reversal of stimulus-reinforcement associations, and thus the correction of behavioural responses when these are no longer appropriate because previous reinforcement contingencies change. The information which reaches the orbitofrontal cortex for these functions includes information about faces, and damage to the orbitofrontal cortex can impair face expression identification. This evidence thus shows that the orbitofrontal cortex is involved in decoding some primary reinforcers such as taste; in learning and reversing associations of visual and other stimuli to these primary reinforcers; and plays an executive function in controlling and correcting reward-related and punishment-related behaviour, and thus in emotion.

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

1. INTRODUCTION

The prefrontal cortex is the cortex that receives projections from the mediodorsal nucleus of the thalamus and is situated in front of the motor and premotor cortices (Areas 4 and 6) in the frontal lobe. Based on the divisions of the mediodorsal nucleus, the prefrontal cortex may be divided into three main regions (Fuster 1989). First, the magnocellular, medial, part of the mediodorsal nucleus projects to the orbital (ventral) surface of the prefrontal cortex (which includes Areas 13 and 12). It is called the orbitofrontal cortex, and receives information from the ventral or object-processing visual stream, and taste, olfactory and somatosensory inputs. Second, the parvocellular, lateral, part of the mediodorsal nucleus projects to the dorsolateral prefrontal cortex. This part of the prefrontal cortex receives inputs from the parietal cortex, and is involved in tasks such as spatial short-term memory tasks (Rosenkilde 1979; Fuster 1989). Third, the pars paralamellaris (most lateral) part of the mediodorsal nucleus projects to the frontal eye fields (Area 8) in the anterior bank of the arcuate sulcus.

The orbitofrontal cortex is considered in this paper. The cortex on the orbital surface of the frontal lobe includes Area 13 caudally, and Area 14 medially, and the cortex on the inferior convexity includes Area 12 caudally and Area 11 anteriorly (see figure 1 and Carmichael & Price 1994; Petrides & Pandya 1995). This brain region is poorly developed in rodents, but well developed in primates including humans. To understand the function of this brain region in humans, the majority of the studies described were therefore performed with macaques or with humans.

2. CONNECTIONS

Rolls *et al.* (1990) discovered a taste area in the lateral part of the orbitofrontal cortex, and showed that this was the secondary taste cortex in that it receives a major projection from the primary taste cortex (Baylis *et al.* 1994). More medially, there is an olfactory area (Rolls & Baylis 1994). Anatomically, there are direct connections from the primary olfactory cortex, pyriform cortex, to area 13a of the posterior orbitofrontal cortex, which in turn has onward projections to a middle part of the orbitofrontal cortex (area 11) (Price *et al.* 1991; Morecraft *et al.* 1992; Barbas 1993; Carmichael *et al.* 1994) (see figures 1 and 2). Visual inputs reach the orbitofrontal cortex directly from the inferior temporal cortex, the cortex in the superior temporal sulcus, and the temporal pole (Jones & Powell 1970; Barbas 1988, 1993; Petrides & Pandya 1988; Seltzer & Pandya 1989; Barbas & Pandya 1989; Morecraft *et al.* 1992; Barbas 1995). There are corresponding auditory inputs (Barbas 1988, 1993), and somatosensory inputs from somatosensory cortical areas 1, 2 and SII in the frontal and pericentral operculum, and from the insula (Barbas 1988; Preuss & Goldman-Rakic 1989). The caudal orbitofrontal cortex receives strong inputs from the amygdala (e.g. Price *et al.* 1991). The orbitofrontal cortex also receives inputs via the mediodorsal nucleus of the thalamus, pars magnocellularis, which itself receives afferents from temporal lobe structures such as the prepyriform (olfactory) cortex, amygdala and inferior temporal cortex (Nauta 1972; Krettek & Price 1974, 1977). The orbitofrontal cortex projects back to temporal lobe areas such as the inferior temporal cortex, and, in

© 1996 The Royal Society

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

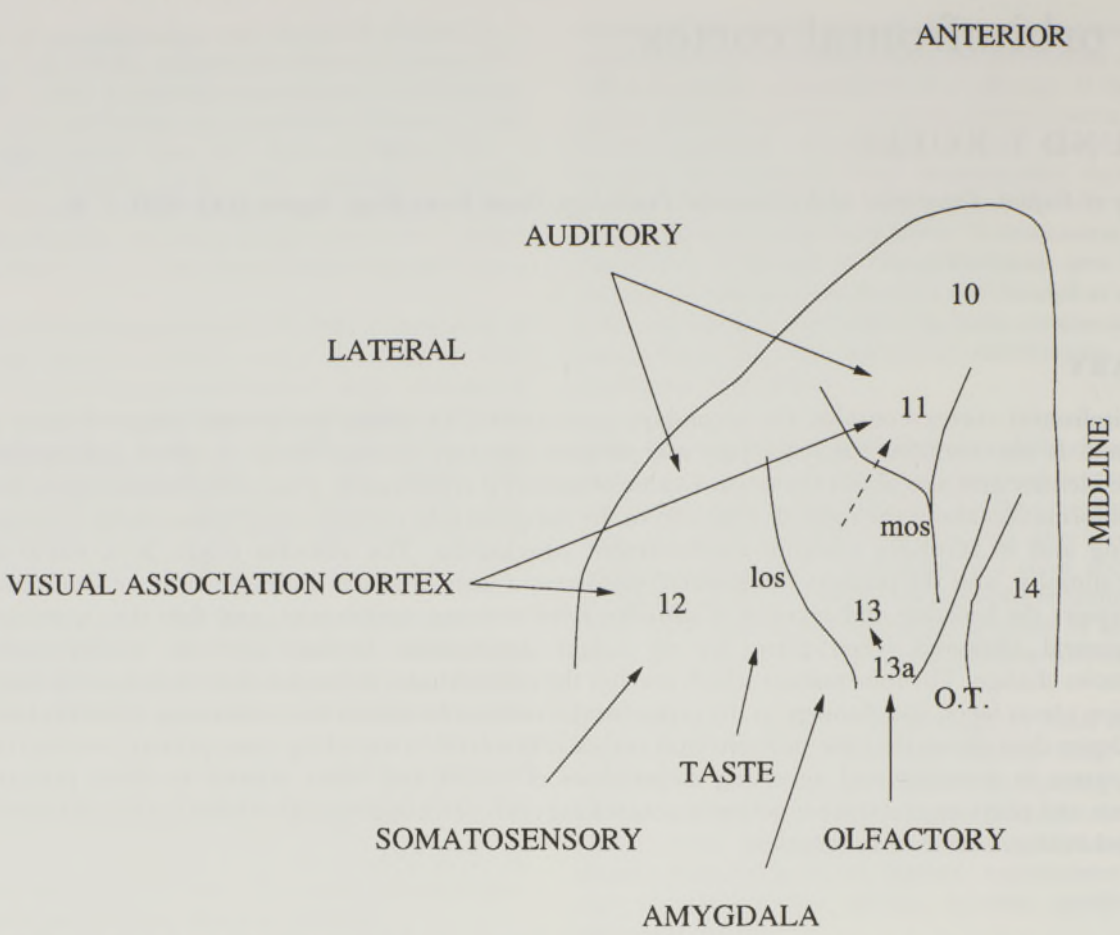

Figure 1. Ventral view of the macaque orbitofrontal cortex. The midline is on the left of the diagram, and the inferior convexity is laterally, on the right. Subdivisions (after Barbas & Pandya 1989) and some afferents to the orbitofrontal cortex are shown. mos, medial orbital sulcus; los, lateral orbital sulcus.

addition, to the entorhinal cortex (or 'gateway to the hippocampus') and cingulate cortex (Nauta 1964; Insausti *et al.* 1987). The orbitofrontal cortex also projects to the preoptic region and lateral hypothalamus, to the ventral tegmental area (Nauta 1964; Johnson *et al.* 1968), and to the head of the caudate nucleus (Kemp & Powell 1970). Reviews of the cytoarchitecture and connections of the orbitofrontal cortex are provided by Carmichael & Price (1994), Barbas (1995), Petrides & Pandya (1995) and Pandya & Yeterian (1996, this issue).

3. EFFECTS OF LESIONS OF THE ORBITOFRONTAL CORTEX

Macaques with lesions of the orbitofrontal cortex are impaired at tasks which involve learning about which stimuli are rewarding and which are not, and especially in altering behaviour when reinforcement contingencies change. The monkeys may respond when responses are inappropriate for example they are no longer rewarded, or may respond to a non-rewarded stimulus. For example, monkeys with orbitofrontal damage are impaired on go/no go task performance, in that they go on the no go trials (Iversen & Mishkin, 1970), in an object reversal task in that they respond to the object which was formerly rewarded with food, and in extinction in that they continue to respond to an object which is no longer rewarded (Butter 1969; Jones & Mishkin 1972). There is some evidence for dissociation of function within the orbitofrontal cortex, in that lesions to the inferior convexity produce the go/no go and object reversal deficits, whereas damage to the caudal orbitofrontal cortex, area 13, produces the extinction deficit (Rosenkilde 1979).

Lesions more laterally, in for example the inferior convexity, can influence tasks in which objects must be remembered for short periods, for example delayed matching to sample and delayed matching to non-sample tasks (Passingham 1975; Mishkin & Manning 1978; Kowalska *et al.* 1991), and neurons in this region may help to implement this visual object short term memory by holding the representation active during the delay period (Rosenkilde *et al.* 1981; Wilson *et al.* 1993). Whether this inferior convexity area is specifically involved in a short term object memory is not yet clear, and a medial part of the frontal cortex may also contribute to this function (Kowalska *et al.* 1991). It should be noted that this short term memory system for objects (which receives inputs from the temporal lobe visual cortical areas in which objects are represented) is different to the short term memory system in the dorsolateral part of the prefrontal cortex, which is concerned with spatial short term memories, consistent with its inputs from the parietal cortex (see e.g. Williams *et al.* 1993).

Damage to the caudal orbitofrontal cortex in the monkey also produces emotional changes (e.g. decreased aggression to humans and to stimuli such as a

snake and a doll), and a reduced tendency to reject foods such as meat (Butter *et al.* 1969; Butter *et al.* 1970; Butler & Snyder, 1972) or to display the normal preference ranking for different foods (Baylis & Gaffan 1991). In the human, euphoria, irresponsibility, and lack of affect can follow frontal lobe damage (see Kolb & Whishaw 1990; Damasio 1994), particularly orbitofrontal damage (Rolls *et al.* 1994).

4. NEUROPHYSIOLOGY OF THE ORBITOFRONTAL CORTEX

(a) *Taste*

One of the recent discoveries that has helped us to understand the functions of the orbitofrontal cortex in behaviour is that it contains a major cortical representation of taste (see Rolls 1989, 1995*a*; cf figure 2). Given that taste can act as a primary reinforcer, that is, without learning as a reward or punishment, we now have the start for a fundamental understanding of the function of the orbitofrontal cortex in stimulus-reinforcement association learning. We know how one class of primary reinforcers reaches and is represented in the orbitofrontal cortex. A representation of primary reinforcers is essential for a system that is involved in learning associations between previously neutral stimuli and primary reinforcers, for example between the sight of an object, and its taste.

The representation (shown by analysing the responses of single neurons in macaques) of taste in the orbitofrontal cortex includes robust representations of the prototypical tastes sweet, salt, bitter and sour (Rolls *et al.* 1990), but also separate representations of the taste of water (Rolls *et al.* 1990), of protein or umami as exemplified by monosodium glutamate (Baylis & Rolls 1991) and inosine monophosphate (Rolls *et al.* 1996*a*), and of astringency as exemplified by tannic acid (Critchley & Rolls 1996*c*).

The nature of the representation of taste in the orbitofrontal cortex is that the reward value of the taste is represented. The evidence for this is that the responses of orbitofrontal taste neurons are modulated by hunger (as is the reward value or palatability of a taste). In particular, it has been shown that orbitofrontal cortex taste neurons stop responding to the taste of a food with which the monkey is fed to satiety (Rolls *et al.* 1989). In contrast, the representation of taste in the primary taste cortex (Scott *et al.* 1986; Yaxley *et al.* 1990) is not modulated by hunger (Rolls *et al.* 1988; Yaxley *et al.* 1988). Thus in the primary taste cortex, the reward value of taste is not represented, and instead the identity of the taste is represented. Additional evidence that the reward value of food is represented in the orbitofrontal cortex is that monkeys work for electrical stimulation of this brain region if they are hungry, but not if they are satiated (Mora *et al.* 1979). Further, neurons in the orbitofrontal cortex are activated from many brain-stimulation reward sites (Mora *et al.* 1980; Rolls *et al.* 1980). Thus there is clear evidence that it is the reward value of taste that is represented in the orbitofrontal cortex.

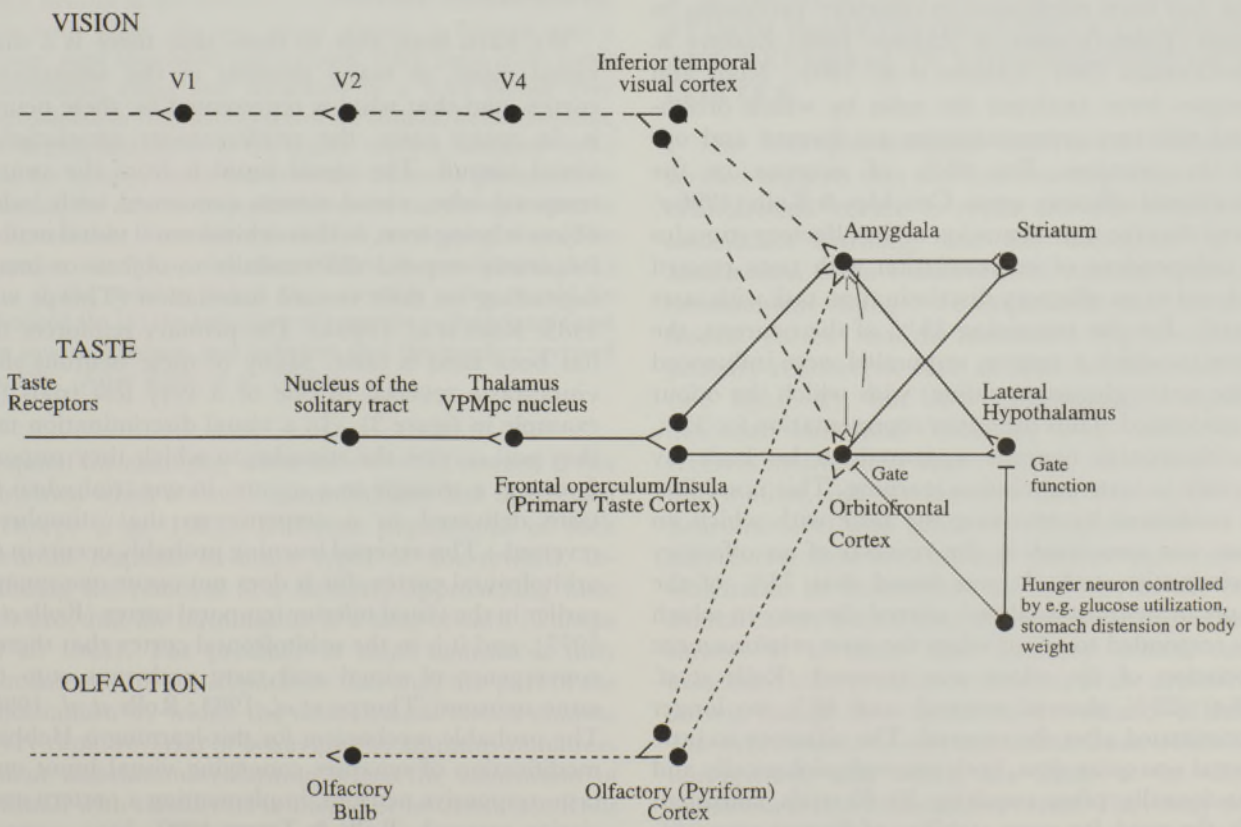

Figure 2. Schematic diagram showing some of the gustatory, olfactory, and visual pathways to the orbitofrontal cortex, and some of the outputs of the orbitofrontal cortex. The secondary taste cortex, and the secondary olfactory cortex, are within the orbitofrontal cortex. V1 – primary visual cortex. V4 – visual cortical area V4.

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

The secondary taste cortex is in the caudolateral part of the orbitofrontal cortex, as defined anatomically (Baylis *et al.* 1994). This region projects on to other regions in the orbitofrontal cortex (Baylis *et al.* 1994), and neurons with taste responses (in what can be considered as a tertiary gustatory cortical area) can be found in many regions of the orbitofrontal cortex (see Rolls *et al.* 1990; Rolls & Baylis 1994; Rolls *et al.* 1996*a*).

(b) *Convergence of taste and olfactory inputs in the orbitofrontal cortex: the representation of flavour*

In these further parts of the orbitofrontal cortex, not only unimodal taste neurons, but also unimodal olfactory neurons are found. In addition some single neurons respond to both gustatory and olfactory stimuli, often with correspondence between the two modalities (Rolls & Baylis 1994; cf figure 2). It is probably here in the orbitofrontal cortex of primates that these two modalities converge to produce the representation of flavour (Rolls & Baylis 1994). Evidence will soon be described that indicates that these representations are built by olfactory-gustatory association learning, an example of stimulus-reinforcement association learning.

(c) *An olfactory representation in the orbitofrontal cortex*

Takagi, Tanabe and colleagues (see Takagi 1991) described single neurons in the macaque orbitofrontal cortex that were activated by odours. A ventral frontal region has been implicated in olfactory processing in humans (Jones-Gotman & Zatorre 1988; Zatorre & Jones-Gotman 1991; Zatorre *et al.* 1992). Rolls and colleagues have analysed the rules by which orbitofrontal olfactory representations are formed and operate in primates. For 65% of neurons in the orbitofrontal olfactory areas, Critchley & Rolls (1996*a*) showed that the representation of the olfactory stimulus was independent of its association with taste reward (analysed in an olfactory discrimination task with taste reward). For the remaining 35% of the neurons, the odours to which a neuron responded were influenced by the taste (glucose or saline) with which the odour was associated. Thus the odour representation for 35% of orbitofrontal neurons appeared to be built by olfactory to taste association learning. This possibility was confirmed by reversing the taste with which an odour was associated in the reversal of an olfactory discrimination task. It was found that 73% of the sample of neurons analysed altered the way in which they responded to odour when the taste reinforcement association of the odour was reversed (Rolls *et al.* 1996*a*) (25% showed reversal, and 48% no longer discriminated after the reversal. The olfactory to taste reversal was quite slow, both neurophysiologically and behaviourally, often requiring 20–80 trials, consistent with the need for some stability of flavour representations. The relatively high proportion of neurons with modification of responsiveness by taste association in the set of neurons in this experiment was probably related to the fact that the neurons were preselected to show differential responses to the odours associated with different tastes in the olfactory discrimination task.) Thus the rule according to which the orbitofrontal olfactory representation was formed was for some neurons by association learning with taste.

To analyse the nature of the olfactory representation in the orbitofrontal cortex, Critchley & Rolls (1996*b*) measured the responses of olfactory neurons that responded to food while they fed the monkey to satiety. They found that the majority of orbitofrontal olfactory neurons decreased their responses to the odour of the food with which the monkey was fed to satiety. Thus for these neurons, the reward value of the odour is what is represented in the orbitofrontal cortex. We do not yet know whether this is the first stage of processing at which reward value is represented in the olfactory system (although in rodents the influence of reward association learning appears to be present in some neurons in the pyriform cortex – Schoenbaum & Eichenbaum 1995).

Although individual neurons do not encode large amounts of information about which of 7–9 odours has been presented, we have shown that the information does increase linearly with the number of neurons in the sample (Rolls *et al.* 1996*b*). This ensemble encoding does result in useful amounts of information about which odour has been presented being provided by orbitofrontal olfactory neurons.

(d) *Visual inputs to the orbitofrontal cortex, and visual stimulus – reinforcement association learning and reversal*

We have been able to show that there is a major visual input to many neurons in the orbitofrontal cortex, and that what is represented by these neurons is, in many cases, the reinforcement association of visual stimuli. The visual input is from the ventral, temporal lobe, visual stream concerned with 'what' object is being seen, in that orbitofrontal visual neurons frequently respond differentially to objects or images depending on their reward association (Thorpe *et al.* 1983; Rolls *et al.* 1996*a*). The primary reinforcer that has been used is taste. Many of these neurons show visual-taste reversal in one or a very few trials (see example in figure 3). (In a visual discrimination task, they will reverse the stimulus to which they respond, from e.g. a triangle to a square, in one trial when the taste delivered to a response to that stimulus is reversed.) This reversal learning probably occurs in the orbitofrontal cortex, for it does not occur one synapse earlier in the visual inferior temporal cortex (Rolls *et al.* 1977), and it is in the orbitofrontal cortex that there is convergence of visual and taste pathways onto the same neurons (Thorpe *et al.* 1983; Rolls *et al.* 1996). The probable mechanism for this learning is Hebbian modification of synapses conveying visual input onto taste-responsive neurons, implementing a pattern association network (Rolls & Treves 1990).

In addition to these neurons that encode the reward association of visual stimuli, other neurons in the orbitofrontal cortex detect non-reward, in that they

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

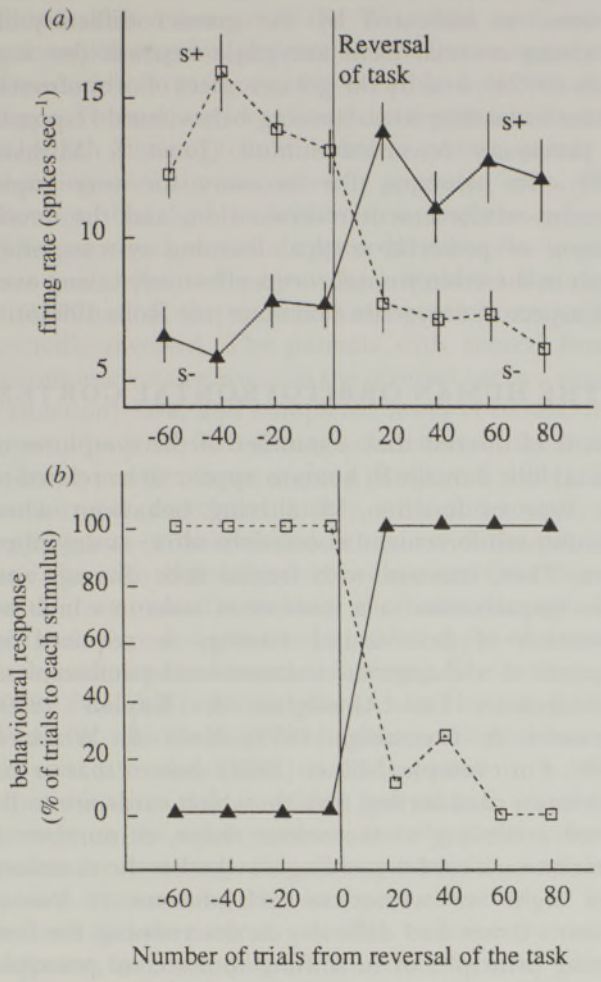

Figure 3. Visual discrimination reversal of the responses of a single neuron in the macaque orbitofrontal cortex when the taste with which the two visual stimuli (a triangle and a square) were associated was reversed. Each point is the mean poststimulus firing rate measured in a 0.5 s period over approximately 10 trials to each of the stimuli. Before reversal, the neuron fired most to the square when it indicated (S+) that the monkey could lick to obtain a taste of glucose. After reversal, the neuron responded most to the triangle when it indicated that the monkey could lick to obtain glucose. The response was low to the stimuli when they indicated (S−) that if the monkey licked then aversive saline would be obtained. In (*b*) the behavioural response to the triangle and the square is shown, and indicates that the monkey reversed rapidly. (After Rolls, *et al.* 1996*a*).

respond for example, when an expected reward is not obtained when a visual discrimination task is reversed (Thorpe *et al.* 1983). Different populations of such neurons respond to other types of non-reward, including the removal of a formerly approaching taste reward, and the termination of a taste reward (Thorpe *et al.* 1983). The presence of these neurons is fully consistent with the hypothesis that they are part of the mechanism by which the orbitofrontal cortex enables very rapid reversal of behaviour by stimulus-reinforcement association relearning when the association of stimuli with reinforcers is altered or reversed (see Rolls 1989*a*, 1990). Different orbitofrontal cortex neurons respond to different types of non-reward (Thorpe *et al.* 1983), potentially enabling task or context-specific reversal to occur.

Another type of information represented in the orbitofrontal cortex is information about faces. There is a population of orbitofrontal neurons which respond in many ways similar to those in the temporal cortical visual areas (see Rolls 1984*a*, 1992*a*, 1994*a*, 1995*b*, 1996 for a description of their properties). The orbitofrontal face responsive neurons, first observed by Thorpe *et al.* (1983), then by Rolls *et al.* (1983), then by Rolls *et al.* (unpublished data), tend to respond with longer latencies than temporal lobe neurons (140–200 ms typically, compared to 80–100 ms); also convey information about which face is being seen, by having different responses to different faces; and are typically rather harder to activate strongly than temporal cortical face-selective neurons, in that many of them respond much better to real faces than to two-dimensional images of faces on a video monitor (see Rolls & Baylis, 1986). Some of the orbitofrontal face-selective neurons are responsive to face gesture or movement. The findings are consistent with the likelihood that these neurons are activated via the inputs from the temporal cortical visual areas in which face-selective neurons are found (see figure 2). The significance of the neurons is likely to be related to the fact that faces convey information that is important in social reinforcement, both by conveying face expression (see Hasselmo *et al.* 1989), which can indicate reinforcement; and by encoding information about which individual is present, also important in evaluating and utilising reinforcing inputs in social situations.

5. A NEUROPHYSIOLOGICAL BASIS FOR STIMULUS-REINFORCEMENT LEARNING AND REVERSAL IN THE ORBITOFRONTAL CORTEX

The neurophysiological and lesion evidence described suggests that one function implemented by the orbitofrontal cortex is rapid stimulus-reinforcement association learning, and the correction of these associations when reinforcement contingencies in the environment change. To implement this, the orbitofrontal cortex has the necessary representation of primary reinforcers, such as taste, as described above (see figure 2), but also somatosensory inputs, as described elsewhere (Critchley *et al.* 1993). It also receives information about objects, for example visual information, and can associate this very rapidly at the neuronal level with primary reinforcers such as taste, and reverse these associations. Another type of stimulus which can be conditioned in this way in the orbitofrontal cortex is olfactory, although here the learning is slower. It is likely that auditory stimuli can be associated with primary reinforcers in the orbitofrontal cortex, though there is less direct evidence of this yet. The orbitofrontal cortex also has neurons which detect non-reward, and which are likely to be used in behavioural extinction and reversal. They may do this not only by helping to reset the reinforcement association of neurons in the orbitofrontal cortex, but also by sending a signal to the striatum which could be routed by the striatum to produce appropriate beha-

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

viours for non-reward (Rolls & Johnstone 1992; Williams *et al.* 1993; Rolls 1994*b*). Indeed, it is via this route, the striatal, that the orbitofrontal cortex may directly influence behaviour when the orbitofrontal cortex is decoding reinforcement contingencies in the environment, and is altering behaviour in response to altering reinforcement contingencies. Some of the evidence for this is that neurons which reflect these orbitofrontal neuronal responses are found in the ventral part of the head of the caudate nucleus and the ventral striatum, which receive from the orbitofrontal cortex (Rolls *et al.* 1983*b*; Williams *et al.* 1993); and lesions of the ventral part of the head of the caudate nucleus impair visual discrimination reversal (Divac *et al.* 1967).

Decoding the reinforcement value of stimuli, which involves for previously neutral (e.g. visual) stimuli learning their association with a primary reinforcer, often rapidly, and which may involve not only rapid learning but also rapid relearning and alteration of responses when reinforcement contingencies change, is then a function proposed for the orbitofrontal cortex. This way of producing behavioural responses would be important in for example motivational and emotional behaviour. It would be important for example in motivational behaviour such as feeding and drinking by enabling primates to learn rapidly about the food reinforcement to be expected from visual stimuli (see Rolls 1994*c*). This is important, for primates frequently eat more than 100 varieties of food; vision by visual-taste association learning can be used to identify when foods are ripe; and during the course of a meal, the pleasantness of the sight of a food eaten in the meal decreases in a sensory-specific way (Rolls *et al.* 1983*a*), a function that is probably implemented by the sensory-specific satiety-related responses of orbitofrontal visual neurons (Critchley & Rolls 1996*b*).

With respect to emotional behaviour, decoding and rapidly readjusting the reinforcement value of visual signals is likely to be crucial, for emotions can be described as responses elicited by reinforcing signals (For the purposes of this paper, a positive reinforcer or reward can be defined as a stimulus which the animal will work to obtain, and a negative reinforcer or punishment as a stimulus that an animal will work to avoid or escape (see further Rolls 1986*a*, *b*, 1990, 1995*b*). The ability to perform this learning very rapidly is probably very important in social situations in primates, in which reinforcing stimuli are continually being exchanged, and the reinforcement value of these must be continually updated (relearned), based on the actual reinforcers received and given. Although the operation of reinforcers such as taste, smell, and faces are best understood in terms of orbitofrontal cortex operation, there are tactile inputs that are likely to be concerned with reward evaluation, and in humans the rewards processed in the orbitofrontal cortex include quite general rewards such as working for 'points', as will be described shortly.

Although the amygdala is concerned with some of the same functions as the orbitofrontal cortex, and receives similar inputs (see figure 2), there is evidence that it may function less effectively in the very rapid learning and reversal of stimulus-reinforcement associations, as indicated by the greater difficulty in obtaining reversal from amygdala neurons (see e.g. Rolls 1992*b*), and by the greater effect of orbitofrontal lesions in leading to continuing behavioural responses to previously rewarded stimuli (Jones & Mishkin 1972). In primates, the necessity for very rapid stimulus-reinforcement re-evaluation, and the development of powerful cortical learning systems, may result in the orbitofrontal cortex effectively taking over this aspect of amygdala functions (see Rolls 1992*b*).

6. THE HUMAN ORBITOFRONTAL CORTEX

It is of interest that a number of the symptoms of frontal lobe damage in humans appear to be related to this type of function, of altering behaviour when stimulus-reinforcement associations alter, as described next. Thus, humans with frontal lobe damage can show impairments in a number of tasks in which an alteration of behavioural strategy is required in response to a change in environmental reinforcement contingencies (see Goodglass & Kaplan 1979; Jouandet & Gazzaniga 1979; Kolb & Whishaw 1990). For example, Milner (1963) showed that in the Wisconsin card sorting task (in which cards are to be sorted according to the colour shape, or number of items on each card depending on whether the examiner says 'right' or 'wrong' to each placement), frontal patients either had difficulty in determining the first sorting principle, or in shifting to a second principle when required to. Also, in stylus mazes, frontal patients have difficulty in changing direction when a sound indicates that the correct path has been left (see Milner 1982). It is of interest that, in both types of test, frontal patients may be able to verbalize the correct rules, yet may be unable to correct their behavioural sets or strategies appropriately. Some of the personality changes that can follow frontal lobe damage may be related to a similar type of dysfunction. For example, the euphoria, irresponsibility, lack of affect, and lack of concern for the present or future which can follow frontal lobe damage (see Hecaen & Albert 1978) may also be related to a dysfunction in altering behaviour appropriately in response to a change in reinforcement contingencies. Indeed, in so far as the orbitofrontal cortex is involved in the disconnection of stimulus reinforcement associations, and such associations are important in learned emotional responses (see above), then it follows that the orbitofrontal cortex is involved in emotional responses by correcting stimulus-reinforcement associations when they become inappropriate.

These hypotheses, and the role in particular of the orbitofrontal cortex in human behaviour, have been investigated in recent studies in humans with damage to the ventral parts of the frontal lobe. (The description ventral is given to indicate that there was pathology in the orbitofrontal or related parts of the frontal lobe, and not in the more dorso-lateral parts of the frontal lobe.) A task which was directed at assessing the rapid alteration of stimulus-reinforcement associations was used, because the findings above indicate that the

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

orbitofrontal cortex is involved in this type of learning. This was used instead of the Wisconsin card sorting task, which requires patients to shift from category (or dimension) to category, for example from colour to shape. The task used was visual discrimination reversal, in which patients could learn to obtain points by touching one stimulus when it appeared on a video monitor, but had to withhold a response when a different visual stimulus appeared, otherwise a point was lost. After the subjects had acquired the visual discrimination, the reinforcement contingencies unexpectedly reversed. The patients with ventral frontal lesions made more errors in the reversal (or in a similar extinction) task, and completed fewer reversals, than control patients with damage elsewhere in the frontal lobes or in other brain regions (Rolls *et al.* 1994). The impairment correlated highly with the socially inappropriate or disinhibited behaviour of the patients, and also with their subjective evaluation of the changes in their emotional state since the brain damage. The patients were not impaired at other types of memory task, such as paired associate learning. The findings are being extended in current research by Hornak, Rolls & Wade in which visual discrimination acquisition and reversal are also found to be impaired in a visual discrimination task in which two stimuli are always present on the video monitor, and the patient obtains points by touching the correct stimulus, and loses points by touching the incorrect stimulus. It is of interest that the patients can often verbalize the correct response, yet commit the incorrect action. This is consistent with the hypothesis that the orbitofrontal cortex is normally involved in executing behaviour when the behaviour is performed by evaluating the reinforcement associations of environmental stimuli (see below). The orbitofrontal cortex appears to be involved in this in both humans and non-human primates, when the learning must be performed rapidly, in for example acquisition, and during reversal.

To investigate the possible significance of face-related inputs to orbitofrontal visual neurons described above, we also tested the responses of these patients to faces. We included tests of face (and also voice) expression decoding, because these are ways in which the reinforcing quality of individuals is often indicated. Impairments in the identification of facial and vocal emotional expression were demonstrated in a group of patients with ventral frontal lobe damage who had socially inappropriate behaviour (Hornak *et al.* 1996). The expression identification impairments could occur independently of perceptual impairments in facial recognition, voice discrimination, or environmental sound recognition. The face and voice expression problems did not necessarily occur together in the same patients, providing an indication of separate processing. Poor performance on both expression tests was correlated with the degree of alteration of emotional experience reported by the patients. There was also a strong positive correlation between the degree of altered emotional experience and the severity of the behavioural problems (e.g. disinhibition) found in these patients. A comparison group of patients with brain damage outside the ventral frontal lobe region, without these behavioural problems, was unimpaired on the face expression identification test, was significantly less impaired at vocal expression identification, and reported little subjective emotional change (Hornak *et al.* 1996). In current studies, these findings are being extended, and it is being found that patients with face expression decoding problems do not necessarily have impairments at visual discrimination reversal, and vice versa. This is consistent with some topography in the orbitofrontal cortex (see e.g. Rolls & Baylis 1994).

7. EXECUTIVE FUNCTIONS OF THE ORBITOFRONTAL CORTEX

The research described indicates that the orbitofrontal cortex is involved in the execution of behavioural responses when these are computed by reward or punishment association learning, a function for which the orbitofrontal cortex is specialised, in terms of representations of primary (unlearned) reinforcers, and in rapidly learning and readjusting associations of stimuli with these primary reinforcers. The fact that patients with ventral frontal lesions often can express verbally what the correct responses should be, yet cannot follow what previously obtained rewards and punishments indicate is appropriate behaviour, is an indication that when primates (including humans) normally execute behavioural responses on the basis of reinforcement evaluation, they do so using the orbitofrontal cortex. Eliciting behaviour on the basis of rewards and punishments obtained previously in similar situations is of course a simple and adaptive way to control behavioural responses that has been studied and accepted for very many years (see e.g. the history of psychology; and in terms of brain mechanisms, see e.g. Rolls 1975, 1986*a*, *b*, 1990, 1994*b*, 1995*b*), and has been recently emphasized by Damasio (1994). The particular utility of one of the alternative routes to behaviour (there are of course many routes to behaviour) made possible by language is that this enables long-term planning, where the plan involves many syntactic arrangements of symbols (e.g. many if... then statements). It is suggested that when this linguistic (in terms of syntactic manipulation of symbols) system needs correction, being able to think about the plans (higher order thoughts), enables the plans to be corrected, and that this process is closely related to explicit, conscious, processing (Rolls 1995*b*, 1997; see also Rosenthal 1993). It follows that the functions performed by the orbitofrontal cortex need not be performed with explicit (conscious) processing, but can be performed with implicit processing. It is in this way that the orbitofrontal cortex is suggested to be involved in some, but certainly not all, types of executive function.

In that the orbitofrontal cortex may retain as a result of synaptic modification in a pattern associator (see Rolls & Treves 1990) the most recent reinforcement association for large numbers of different stimuli, it could perhaps be fitted into a view that the frontal cortical areas are in general concerned with different

types of working memory. However, the term working memory is normally used in neurophysiology to refer to a memory in which the memoranda are held in the memory by continuing neuronal activity, as in an autoassociator or attractor network (see e.g. Treves & Rolls 1991). It should be realised that although there may be a functional similarity between such a working memory and the ability of the orbitofrontal cortex to retain the most recent reinforcement association of many stimuli, the implementations are very different. The different implementations do in fact have strong functional consequences: it is difficult to retain more than a few items active in an autoassociative memory, and hence in practice individual items are retained typically only for short periods in such working memories; whereas in pattern associators, because synaptic modification has taken place, the last reinforcement association of a very large number of stimuli can be stored for long periods, and recalled whenever each stimulus is seen again in the future, without any neuronal firing to hold the representation active (see e.g. Rolls & Treves 1997).

It is perhaps useful to note how the orbitofrontal cortex may link to output systems to control behaviour, for the occasions when the orbitofrontal cortex does control behaviour. Rolls has proposed elsewhere (Rolls 1984*b*, 1994*b*; Rolls & Johnstone 1992) the outline of a theory of striatal function, according to which all areas of the cerebral cortex gain access to the striatum, compete within the striatum and rest of the basal ganglia system for behavioural output depending on how strongly each part of the cerebral cortex is calling for output, and the striatum maps (as a result of slow previous habit or stimulus-response learning) each particular type of input to the striatum to the appropriate behavioural output (implemented via the return basal ganglia connections to premotor/prefrontal parts of the cerebral cortex). This is one of the ways in which reinforcing stimuli can exert their influence relatively directly on behavioural output. The importance of this route is attested to by the fact that restricted striatal lesions impair functions implemented by the part of the cortex which projects to the lesioned part of the striatum (see Rolls 1984*b*, 1994*b*; Rolls & Johnstone 1992). This hypothesis is very different from that of Damasio (1994), who has effectively tried to resurrect a weakened version of the James-Lange theory of emotion from the last century, by arguing with his somatic marker hypothesis that after reinforcers have been evaluated, a bodily response ('somatic marker') normally occurs, then this leads to a bodily feeling, which in turn is appreciated by the organism to then make a contribution to the decision-making process. (In the James-Lange theory, it was emotional feelings that depend on peripheral feedback; for Damasio, it is the decision of which behavioural response to make that is normally influenced by the peripheral feedback.) The James-Lange theory has a number of major weaknesses, including the evidence that inactivation of peripheral feedback does little to abolish feelings or behaviour to emotion-provoking (reinforcing) stimuli (see Grossman (1967) for an extensive review of this literature; Schachter & Singer (1962), who could alter the magnitude but not the quality of experienced emotion by artificially inducing peripheral feedback; and Reisenzein (1983), who produced by pharmacological blockers little reduction of emotion); does not in any case account for the fundamental question analysed here, of how it is that some stimuli produce emotional responses and others do not (see Rolls 1990), that is the decoding of whether a stimulus is associated with reinforcement that must be performed according to both the direct and peripheral hypotheses; and does not take account of the fact that once an information processor has determined that a response should be made or inhibited based on reinforcement association, a function attributed here in part to the orbitofrontal cortex, it would be very inefficient and noisy to place in the execution route a peripheral response, and transducers to attempt to measure that peripheral response, itself a notoriously difficult procedure (see e.g. Grossman 1967). Even if Damasio were to argue that the peripheral somatic marker and its feedback can be bypassed using conditioning of a representation in for example, the somatosensory cortex to the ventral prefrontal command signal, he apparently would still wish to argue that the activity in the somatosensory cortex is important for the emotion to be appreciated or to influence behaviour. (Without this, the somatic marker hypothesis would vanish.) The prediction would apparently be that if an emotional response were produced to a visual stimulus, then this would necessarily involve activity in the somatosensory cortex or other brain region in which the 'somatic marker' would be represented. This prediction could be tested (for example in patients with somatosensory cortex damage), but it seems most unlikely that an emotion produced by a visual reinforcer would require activity in the somatosensory cortex. The alternative view proposed here (and by Rolls 1990) is that where the reinforcement value of the visual stimulus is decoded, namely in the orbitofrontal cortex and the amygdala, is the appropriate part of the brain for outputs to influence behaviour (via e.g. the orbitofrontal-to-striatal connections), and that the orbitofrontal cortex is the likely place where neuronal activity is directly related to the felt emotion (see further Rolls 1997).

The fact that ventral prefrontal lesions block autonomic responses to learned reinforcers (Damasio 1994) (actually known since at least the 1950s, e.g. Elithorn *et al.* 1955 in humans; Grueninger *et al.* 1965 in macaques) is of course consistent with the hypothesis that learned reinforcers elicit autonomic responses via the orbitofrontal cortex and amygdala (see e.g. Rolls 1986*a*, *b*, 1990); but does not prove the hypothesis that behavioural responses elicited by conditioned reinforcers are mediated via peripheral changes, themselves used as 'somatic markers' to determine which response to make. Instead, the much more direct neural route from the orbitofrontal cortex and amygdala to the basal ganglia provides a pathway which is much more efficient, and is directly implicated in producing, the behavioural responses to learned incentives (Rolls 1994*b*; Williams *et al.* 1993; Everitt & Robbins 1992; Divac *et al.* 1967).

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

The author has worked on some of the experiments described here with L. L. Baylis, G. C. Baylis, H. Critchley, M. E. Hasselmo, J. Hornak, C. M. Leonard, F. Mora, D. I. Perrett, T. R. Scott, S. J. Thorpe, E. A. Wakeman and F. A. W. Wilson, and their collaboration is sincerely acknowledged. Some of the research described was supported by the Medical Research Council, PG8513790.

REFERENCES

Barbas, H. 1988 Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* **276**, 313–342.

Barbas, H. 1993 Organization of cortical afferent input to the orbitofrontal area in the rhesus monkey. *Neuroscience* **56**, 841–864.

Barbas, H. 1995 Anatomic basis of cognitive-emotional interactions in the primate prefrontal cortex. *Neurosci. Biobehav. Rev.* **19**, 499–510.

Barbas, H. & Pandya, D. N. 1989 Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* **286**, 353–375.

Baylis, L. L. & Gaffan, D. 1991 Amygdalectomy and ventromedial prefrontal ablation produce similar deficits in food choice and in simple object discrimination learning for an unseen reward. *Exp. Brain Res.* **86**, 617–622.

Baylis, L. L. & Rolls, E. T. 1991 Responses of neurons in the primate taste cortex to glutamate. *Physiol. Behav.* **49**, 973–979.

Baylis, L. L., Rolls, E. T. & Baylis, G. C. 1994 Afferent connections of the orbitofrontal cortex taste area of the primate. *Neuroscience* **64**, 801–812.

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

Butter, C. M., McDonald, J. A. & Snyder, D. R. 1969 Orality, preference behavior, and reinforcement value of non-food objects in monkeys with orbital frontal lesions. *Science* **164**, 1306–1307.

Butter, C. M. & Snyder, D. R. 1972 Alterations in aversive and aggressive behaviors following orbitofrontal lesions in rhesus monkeys. *Acta Neurobiol. Exp.* **32**, 525–565.

Butter, C. M., Snyder, D. R. & McDonald, J. A. 1970 Effects of orbitofrontal lesions on aversive and aggressive behaviors in rhesus monkeys. *J. Comp. Physiol. Psychol.* **72**, 132–144.

Carmichael, S. T., Clugnet, M.-C. & Price, J. L. 1994 Central olfactory connections in the macaque monkey. *J. Comp. Neurol.* **346**, 403–434.

Carmichael, S. T. & Price, J. L. 1994 Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J. Comp. Neurol.* **346**, 366–402.

Critchley, H. D., Rolls, E. T. & Wakeman, E. A. 1993 Orbitofrontal cortex responses to the texture, taste, smell and sight of food. *Appetite* **21**, 170.

Critchley, H. D. & Rolls, E. T. 1996*a* Olfactory neuronal responses in the primate orbitofrontal cortex: analysis in an olfactory discrimination task. *J. Neurophysiol.* **75**, 1659–1672.

Critchley, H. D. & Rolls, E. T. 1996*b* Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *J. Neurophysiol.* **75**, 1673–1686.

Critchley, H. D. & Rolls, E. T. 1996*c* Responses of primate taste cortex neurons to the astringent tastant tannic acid. *Chem. Senses* **21**, 135–145.

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

Divac, I., Rosvold, H. E., Szwarcbart, M. K. 1967 Behavioral effects of selective ablation of the caudate nucleus. *J. Comp. Physiol. Psych.* **63**, 184–190.

Elithorn, A., Piercy, M. F. & Crosskey, M. A. 1955 Prefrontal leucotomy and the anticipation of pain. *J. Neurol. Neurosurg. Psychiat.* **18**, 34–43.

Everitt, B. J. & Robbins, T. W. 1992 Amygdala-ventral striatal interactions and reward-related processes. In *The Amygdala* (ed. J. P. Aggleton), pp. 401–430. Chichester: Wiley.

Fuster, J. M. 1989 *The Prefrontal Cortex*, second edition. New York: Raven Press.

Goodglass, H. & Kaplan, E. 1979 Assessment of cognitive deficit in brain-injured patient. In *Handbook of Behavioral Neurobiology. Vol 2 Neuropsychology* (ed. M. S. Gazzaniga), pp. 3–22. New York: Plenum.

Grossman, S. . 1967 *A Textbook of Physiological Psychology*. New York: Wiley.

Grueninger, W. E., Kimble, D. P., Grueninger, J. & Levine, S. 1965 GSR and corticosteroid response in monkeys with frontal ablations. *Neuropsychologia* **3**, 205–216.

Hasselmo, M. E., Rolls, E. T. & Baylis, G. C. 1989 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**, 203–218.

Hecaen, H. & Albert, M. L. 1978 *Human neuropsychology*. New York: Wiley.

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.

Insausti, R., Amaral, D. G. & Cowan, W. M. 1987 The entorhinal cortex of the monkey. II. Cortical afferents. *J. Comp. Neurol.* **264**, 356–395.

Iversen, S. D. & Mishkin, M. 1970 Perseverative interference in monkey following selective lesions of the inferior prefrontal convexity. *Exptl. Brain Res.* **11**, 376–86.

Johnson, T. N., Rosvold, H. E. & Mishkin, M. 1968 Projections from behaviorally defined sectors of the prefrontal cortex to the basal ganglia, septum and diencephalon of the monkey. *Exptl. Neurol.* **21**, 20–34.

Jones, B. & Mishkin, M. 1972 Limbic lesions and the problem of stimulus-reinforcement associations. *Exptl. Neurol.* **36**, 362–77.

Jones, E. G. & Powell, T. P. S. 1970 An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, **93**, 793–820.

Jones-Gotman, M. & Zatorre, R. J. 1988 Olfactory identification in patients with focal cerebral excision. *Neuropsychologia*, **26**, 387–400.

Jouandet, M. & Gazzaniga, M. S. 1979 The frontal lobes. In *Handbook of Behavioral Neurobiology. Vol 2 Neuropsychology* (ed. M. S. Gazzaniga), pp. 25–59. New York: Plenum.

Kemp, J. M. & Powell, T. P. S. 1970 The cortico-striate projections in the monkey. *Brain*, **93**, 525–546.

Kolb, B. & Whishaw, I. W. 1990 *Fundamentals of human neuropsychology*, 3rd edn. New York: Freeman.

Kowalska, D-M., Bachevalier, J. & Mishkin, M. 191 The role of the inferior prefrontal convexity in performance of delayed nonmatching-to-sample. *Neuropsychologia* **29**, 583–600.

Krettek, J. E. & Price, J. L. 1974 A direct input from the amygdala to the thalamus and the cerebral cortex. *Brain Res.* **67**, 169–174.

Krettek, J. E. & Price, J. L. 1977 The cortical projections of the mediodorsal nucleus and adjacent thalamic nuclei in the rat. *J. Comp. Neurol.* **171**, 157–192.

Milner, B. 1963 Effects of different brain lesions on card sorting. *Arch. Neurol.* **9**, 90–100.

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

Milner, B. 1982 Some cognitive effects of frontal-lobe lesions in man. *Phil. Trans. R. Soc. Lond.* B 298, 211–226.

Mishkin, M. & Manning, F. J. 1978 Non-spatial memory after selective prefrontal lesions in monkeys. *Brain Res.* **143**, 313–324.

Mora, F., Avrith, D. B., Phillips, A. G. & Rolls, E. T. 1979 Effects of satiety on self-stimulation of the orbitofrontal cortex in the monkey. *Neuroscience Letters* **13**, 141–145.

Mora, F., Avrith, D. B. & Rolls, E. T. 1980 An electrophysiological and behavioural study of self-stimulation in the orbitofrontal cortex of the rhesus monkey. *Brain Res. Bull.* **5**, 111–115.

Morecraft, R. J., Geula, C., & Mesulam, M.-M. 1992 Cytoarchitecture and neural afferents of orbitofrontal cortex in the brain of the monkey. *J. Comp. Neurol.* **323**, 341–358.

Nauta, W. J. H. 1964 Some efferent connections of the prefrontal cortex in the monkey. In *The Frontal Granular Cortex and Behavior* (ed. J. M. Warren & K. Akert), pp. 397–407. New York: McGraw Hill.

Nauta, W. J. H. 1972 Neural associations of the frontal cortex. *Acta. Neurobiol. Exp.* **32**, 125–140.

Norgren, R. 1984 Central neural mechanisms of taste. In *Handbook of Physiology – The Nervous System III, Sensory Processes 1.* (I. Darien-Smith, editor; Section Editors, J. Brookhart and V. B. Mountcastle), pp. 1087–1128. Washington, DC: American Physiological Society.

Passingham, R. 1975 Delayed matching after selective prefrontal lesions in monkeys (*Macaca mulatta*). *Brain Res.* **92**, 89–102.

Petrides, M. & Pandya, D. N. 1988 Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J. Comp. Neurol.* **273**, 52–66.

Petrides, M. & Pandya, D. N. 1995 Comparative architectonic analysis of the human and macaque frontal cortex. In *Handbook of neuropsychology* (ed. J. Grafman & F. Boller). Elsevier Science Publishers, Amsterdam.

Preuss, T. M. & Goldman-Rakic, P. S. 1989 Connections of the ventral granular frontal cortex of macaques with perisylvian premotor and somatosensory areas: anatomical evidence for somatic representation in primate frontal association cortex. *J. Comp. Neurol.* **282**, 293–316.

Price, J. L., Carmichael, S. T., Carnes, K. M., Clugnet, M.-C., Kuroda, M. & Ray, J. O. 1991 Olfactory input to the prefrontal cortex. In *Olfaction: a model system for computational neuroscience* (ed. J. L. Davis & H. Eichenbaum), pp. 101–120. Cambridge, MA: MIT Press.

Reisenzein, R. 1983 The Schachter theory of emotion: two decades later. *Psychol. Bull.* **94**, 239–264.

Rolls, E. T. 1975 *The brain and reward.* Oxford: Pergamon.

Rolls, E. T. 1984*a* Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum. Neurobiol.* **3**, 209–222.

Rolls, E. T. 1984b Activity of neurons in different regions of the striatum of the monkey. In *The basal ganglia: structure and function* (ed. J. S. McKenzie, R. E. Kemm & L. N. Wilcox), pp. 467–493. New York: Plenum.

Rolls, E. T. 1986*a* A theory of emotion, and its application to understanding the neural basis of emotion. In *Emotions. Neural and chemical control* (ed. Y. Oomura), pp. 325–344. Tokyo: Japan Scientific Societies Press, and Karger: Basel.

Rolls, E. T. 1986*b* Neural systems involved in emotion in primates. In *Emotion: theory, research, and experience.* Volume 3: *Biological foundations of emotion* (ed. R. Plutchik & H. Kellerman), pp. 125–143. New York: Academic Press.

Rolls, E. T. 1989 Information processing in the taste system of primates. *J. Exptl. Biol.* **146**, 141–164.

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

Rolls, E. T. 1992*a* Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Phil. Trans. R. Soc. Lond.* B **335**, 11–21.

Rolls, E. T. 1992*b* Neurophysiology and functions of the primate amygdala. In *The Amygdala* (ed. J. P. Aggleton), ch. 5, pp. 143–165, New York: Wiley-Liss.

Rolls, E. T. 1994*a* Brain mechanisms for invariant visual recognition and learning. *Behav. Proc.* **33**, 113–138.

Rolls, E. T. 1994*b* Neurophysiology and cognitive functions of the striatum. *Revue Neurologique (Paris)* **150**, 648–660.

Rolls, E. T. 1994*c* Neural processing related to feeding in primates. In *Appetite: neural and behavioural bases* (ed. C. R. Legg & D. A. Booth), ch. 2, pp. 11–53, Oxford: Oxford University Press.

Rolls, E. T. 1995*a* Central taste anatomy and neurophysiology. In *Handbook of olfaction and gustation* (ed. R. L. Doty), ch. 24, pp. 549–573, New York: Dekker.

Rolls, E. T. 1995*b* A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In *The cognitive neurosciences* (ed. M. S. Gazzaniga), ch. 72, pp. 1091–1106. Cambridge, MA: MIT Press.

Rolls, E. T. 1996 A neurophysiological and computational approach to the functions of the temporal lobe cortical visual areas in invariant object recognition. In *Computational and biological mechanisms of visual coding* (ed. L. Harris & M. Jenkin), Cambridge: Cambridge University Press.

Rolls, E. T. 1997 Brain mechanisms involved in perception and memory, and their relation to consciousness. In; *Cognition, computation, and consciousness* (ed. M. Ito, Y. Miyashita & E. T. Rolls), Oxford: Oxford University Press.

Rolls, E. T. & Baylis, G. C. 1986 Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Exp. Brain Res.* **65**, 38–48.

Rolls, E. T. & Baylis, L. L. 1994 Gustatory, olfactory and visual convergence within the primate orbitofrontal cortex. *J. Neurosci.* **14**, 5437–5452.

Rolls, E. T., Burton, M. J. & Mora, F. 1980 Neurophysiological analysis of brain-stimulation reward in the monkey. *Brain Res.* **194**, 339–357.

Rolls, E. T., Critchley, H., Mason, R. & Wakeman, E. A. 1996*a* Orbitofrontal cortex neurons: role in olfactory and visual association learning. *J. Neurophysiol.* **75**, 1970–1981.

Rolls, E. T., Critchley, H. D. & Treves, A. 1996*b* The representation of olfactory information in the primate orbitofrontal cortex. *J. Neurophysiol.* **75**, 1982–1996.

Rolls, E. T., Critchley, H., Wakeman, E. A. & Mason, R. 1996*c* Responses of neurons in the primate taste cortex to the glutamate ion and to inosine 5′-monophosphate. *Physiol. Behav.* **59**, 991–1000.

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. Psychiat.* **57**, 1518–1524.

Rolls, E. T. & Johnstone, S. 1992 Neurophysiological analysis of striatal function. In *Neuropsychological disorders associated with subcortical lesions* (ed. G. Vallar, S. F. Cappa & C. W. Wallesch), ch. 3, pp. 61–97, Oxford: Oxford University Press.

Rolls, E. T., Judge, S. J. & Sanghera, M. 1977 Activity of neurons in the inferotemporal cortex of the alert monkey. *Brain Res.* **130**, 229–238.

Rolls, E. T., Rolls, B. J. & Rowe, E. A. 1983*a* Sensory-specific and motivation-specific satiety for the sight and taste of food and water in man. *Physiol. Behav.* **30**, 185–192.

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

Rolls, E. T., Scott, T. R., Sienkiewicz, Z. J. & Yaxley, S. 1988 The responsiveness of neurones in the frontal opercular gustatory cortex of the macaque monkey is independent of hunger. *J. Physiol.* **397**, 1–12.

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 J. Neurosci.* **1**, 53–60.

Rolls, E. T., Thorpe, S. J. & Maddison, S. P. 1983*b* Responses of striatal neurons in the behaving monkey. 1. Head of the caudate nucleus. *Behav. Brain Res.* **7**, 179–210.

Rolls, E. T. & Treves, A. 1990 The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain. *Network* **1**, 407–421.

Rolls, E. T. & Treves, A. 1997 *Neuronal networks and brain function*, Oxford, Oxford University Press.

Rolls, E. T., Yaxley, S. & Sienkiewicz, Z. J. 1990 Gustatory responses of single neurons in the orbitofrontal cortex of the macaque monkey. *J. Neurophysiol.* **64**, 1055–1066.

Rosenkilde, C. E. 1979 Functional heterogeneity of the prefrontal cortex in the monkey: a review. *Behav. Neural Biol.* **25**, 301–345.

Rosenkilde, C. E., Bauer, R. H. & Fuster, J. M. 1981 Single unit activity in ventral prefrontal cortex in behaving monkeys, *Brain Res.* **209**, 375–394.

Rosenthal, D. M. 1993 Thinking that one thinks. In *Consciousness* (ed. M. Davies & G. W. Humphreys), ch. 10, pp. 197–223, Oxford: Blackwell.

Schachter, S. & Singer, J. 1962 Cognitive, social and physiological determinants of emotional state. *Psychol. Rev.* **69**, 378–399.

Schoenbaum, G. & Eichenbaum, H. 1995 Information encoding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *J. Neurophysiol.* **74**, 733–750.

Scott, T. R., Yaxley, S., Sienkiewicz, Z. J. & Rolls, E. T. 1986 Gustatory responses in the frontal opercular cortex of the alert cynomolgus monkey. *J. Neurophysiol.* **56**, 876–890.

Seltzer, B. & Pandya, D. N. 1989 Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *J. Comp. Neurol.* **281**, 97–113.

Takagi, S. F. 1991 Olfactory frontal cortex and multiple olfactory processing in primates. In *Cerebral Cortex. 9* (ed. A. Peters & E. G. Jones), pp. 133–152, New York: Plenum Press.

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

Treves, A. & Rolls, E. T. 1991 What determines the capacity of autoassociative memories in the brain ? *Network* **2**, 371–397.

Williams, G. V., Rolls, E. T., Leonard, C. M. & Stern, C. 1993 Neuronal responses in the ventral striatum of the behaving monkey. *Behav. Brain Res.* **55**, 243–252.

Wilson, F. A. W., Scalaidhe, S. P. O. & Goldman-Rakic, P. S. 1993 Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* **260**, 1955–1958.

Yaxley, S., Rolls, E. T. & Sienkiewicz, Z. J. 1988 The responsiveness of neurones in the insular gustatory cortex of the macaque monkey is independent of hunger. *Physiol. Behav.* **42**, 223–229.

Yaxley, S., Rolls, E. T. & Sienkiewicz, Z. J. 1990 Gustatory responses of single neurons in the insula of the macaque monkey. *J. Neurophysiol.* **63**, 689–700.

Zatorre, R. J. & Jones-Gotman, M. 1991 Human olfactory discrimination after unilateral frontal or temporal lobectomy. *Brain*, **114**, 71–84.

Zatorre, R. J., Jones-Gotman, M., Evans, A. C. & Meyer, E. 1992 Functional localization of human olfactory cortex. *Nature*, **360**, 339–340.

Discussion

B. J. Everitt (*Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K.*). The function that you ascribe to the orbitofrontal cortex, namely association of stimuli with reward value, is a process that many others have ascribed to the amygdala, including interactions with the striatum in terms of access to behavioural output. Would you like to comment on the differences and similarities between the processes occurring within the amygdala and orbitofrontal cortex?

E. T. Rolls. Although the amygdala is concerned with some of the same functions as the orbitofrontal cortex, and receives similar inputs (see figure 2), there is evidence that it may function less effectively in the very rapid learning and reversal of stimulus-reinforcement associations, as indicated by the greater difficulty in obtaining reversal from amygdala neurons (see e.g. Rolls 1992*b*), and by the greater effect of orbitofrontal lesions in leading to continuing behavioural responses to previously rewarded stimuli (Jones & Mishkin 1972). In primates, the necessity for very rapid stimulus-reinforcement re-evaluation, and the development of powerful cortical learning systems, may result in the orbitofrontal cortex being used especially when rapid relearning of stimulus-reinforcement associations is required (see Rolls 1992*b*). The orbitofrontal cortex does develop greatly in primates relative to the amygdala. Indeed, even the taste pathways are connected differently from rodents, with the direct pathways to subcortical structures such as the amygdala and hypothalamus in rodents being replaced by emphasis on connections via the taste thalamus to the primary taste cortex in primates, which is just behind the orbitofrontal cortex, and projects massively to the orbitofrontal cortex (see Norgren 1984; Rolls 1997*b*, 1989, 1995*a*). The increased role of the orbitofrontal cortex in stimulus-reinforcement association learning in primates may be partly related to the fact that primary reinforcers such as taste have a major representation in the primate orbitofrontal cortex. However, it is useful to emphasise that the orbitofrontal cortex is involved in rapid stimulus-reinforcement learning and reversal for a number of different types of reinforcer, in that the patients with impaired stimulus-reinforcement reversal or extinction were working for quite abstract rewards, 'points' awarded to them in the tasks (Rolls *et al.* 1994).

A. Roberts (*Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, U.K.*). An elegant study in monkeys by Iversen and Mishkin in 1970 demonstrated a double dissociation between the effects of lateral and medial regions of orbitofrontal cortex on performance of a visual discrimination reversal task. Whilst a failure to inhibit responding to the previously rewarded stimulus was responsible for the deficit following lesions of lateral orbitofrontal cortex, an impaired ability to learn to respond to the previously non-rewarded stimulus was responsible for the deficit following medial orbitofrontal lesions. Is this functional distinction between the lateral and medial orbitofrontal cortex reflected in the anatomical distribution of the 'unexpected non-reward' neurons that you describe which fire when an expected reward is not obtained following reversal of a visual discrimination.

E. T. Rolls. Taking the studies of Butter (1969) and

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024

Iversen & Mishkin (1970) together, it is clear that macaques with damage to the postero-medial and the lateral orbitofrontal cortex are impaired on a variety of tasks in which responses must not be made to a non-rewarded stimulus in a Go/No Go visual discrimination task; and must not be made to a no-longer rewarded stimulus during extinction or in a visual discrimination reversal. However, a clear and consistent pattern does not emerge from these two studies taken together of exactly what the difference is, if any, between the medial and lateral orbitofrontal cortex in these non-reward tasks. What is quite clear now from the neurophysiology is that the different types of non-reward in different tasks (e.g. extinction in a visual discrimination task; extinction in an ad lib licking task; reversal in a visual discrimination task; and a procedure as simple as withdrawing and removing from sight an expected food reward) each activate different, overlapping, sets of neurons in the orbitofrontal cortex (Thorpe *et al.* 1983). The implication is that when a primate learns about non-reward in one situation, it will not *necessarily* generalize the non-reward to other tasks or situations. In other words, there is some context-specificity of the representation of non-reward in the primate orbitofrontal cortex. Whether there is much topographical separation of the different types of neuron is not clear: they are at least partly intermixed (see Thorpe *et al.* 1983).

It is worth noting that the deficits produced by orbitofrontal cortex lesions in object reversal and extinction do dissociate from effects of lesions in other sites on these and comparison tasks. For example, lesions of the dorsolateral prefrontal cortex and hippocampal region affect spatial reversal much more than object reversal (Butter 1969; Jones & Mishkin 1972); and lesions of the amygdala are interpreted to affect the acquisition of stimulus-reinforcement associations, whereas orbitofrontal lesions impair breaking such stimulus-reinforcement associations (Jones & Mishkin 1972).

Downloaded from https://royalsocietypublishing.org/ on 18 January 2024