Reward representations and reward-related learning in the human brain: insights from neuroimaging

John P O'Doherty[1,2]

This review outlines recent findings from human neuroimaging concerning the role of a highly interconnected network of brain areas including orbital and medial prefrontal cortex, amygdala, striatum and dopaminergic mid-brain in reward processing. Distinct reward-related functions can be attributed to different components of this network. Orbitofrontal cortex is involved in coding stimulus reward value and in concert with the amygdala and ventral striatum is implicated in representing predicted future reward. Such representations can be used to guide action selection for reward, a process that depends, at least in part, on orbital and medial prefrontal cortex as well as dorsal striatum.

Addresses
[1] Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, London WC1N 3BG, UK
[2] Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, USA
e-mail: jdoherty@hss.caltech.edu

Current Opinion in Neurobiology 2004, **14**:769–776

This review comes from a themed issue on
Neurobiology of behaviour
Edited by Alexander Borst and Wolfram Schultz

Available online 10th November 2004

0959-4388/$ – see front matter

DOI 10.1016/j.conb.2004.10.016

Abbreviations

CS	conditioned stimulus
fMRI	functional magnetic resonance imaging
OFC	orbitofrontal cortex
PET	positron emission tomography
UCS	unconditioned stimulus

Introduction

It is axiomatic that most animals including humans have a propensity to seek out rewards and avoid punishments. Central to the organization of such behaviour is the ability to represent the value of rewarding and punishing stimuli, establish predictions of when and where such rewards and punishments will occur and use those predictions to form the basis of decisions that guide behaviour. This review sets out recent advances in understanding the neural substrates of reward processing in the human brain that have arisen from research in functional neuroimaging. The focus is on the role of specific brain structures implicated in reward processing and reward-learning on the basis of extensive research in animals and lesion studies in humans, including the ventromedial prefrontal cortex (encompassing orbital and medial prefrontal regions), amygdala, striatum and dopaminergic midbrain. These regions are highly interconnected and together can be considered as an integrated network.

Representation of stimulus reward value

Role of orbitofrontal cortex

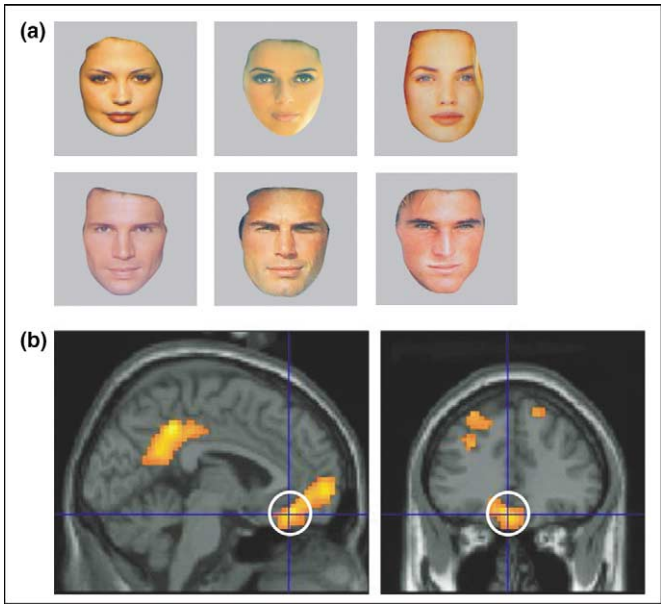
Single unit studies in non-human primates implicate one component of the reward network in particular in coding for stimulus reward value: orbitofrontal cortex (OFC). Neurons in this region respond to a particular taste or odour when an animal is hungry but decrease their firing rate once the animal is satiated and the corresponding food is no longer rewarding [1,2]. Human neuroimaging studies have confirmed a role for human OFC in coding stimulus value from a variety of sensory modalities, including taste [3,4], olfaction [5,6,7••], somatosensory [8], auditory [9] and vision [10–12] as well as for more abstract rewards such as money (Figure 1; [13]). In most of these studies the approach has been to compare OFC activity elicited by an affectively pleasant stimulus with that activity elicited by an affectively neutral stimulus. This leaves open the possibility that such effects are related to differences in sensory properties of the stimuli and not their reward value *per se*.

A different approach makes use of the phenomenon of selective or sensory-specific satiation [14]. This involves scanning hungry subjects during presentation of two food-related stimuli, such as the odour of the corresponding food or else the whole food stimulus itself [15,16]. Subjects are then fed to satiety on one of the corresponding food stimuli, leading to a selective decrease in the reward value of the food eaten, and then scanned again in the satiated state. OFC responses track the reward value of the two foods; activity to the food eaten decreases from pre- to post- satiety but activity to the food not eaten shows no decrease. Given that the experiments involve comparing the same food stimuli before and after satiety, so that the sensory properties are identical but only the reward value changes, these studies provide very strong evidence for a role of human OFC in coding rewarding rather than sensory aspects of a stimulus.

Role of amygdala

Another region implicated in processing stimulus reward value is the amygdala. Although this region has long been known to be involved in appetitive processing in animal literature, early neuroimaging studies tended to focus on its role in responding to aversive stimuli such as fearful

Figure 1



Representation of stimulus–reward value in human orbitofrontal cortex. This is an example from an fMRI study on facial attractiveness. **(a)** Example face stimuli used in the study **(b)**. An area of medial orbitofrontal cortex was found to have enhanced responses to presentation of attractive compared with responses to unattractive faces. (Adapted with permission from [11].)

face expressions or aversive odours [17,18]. Subsequently, evidence has emerged of amygdala responses to pleasant as well as aversive stimuli [3,19].

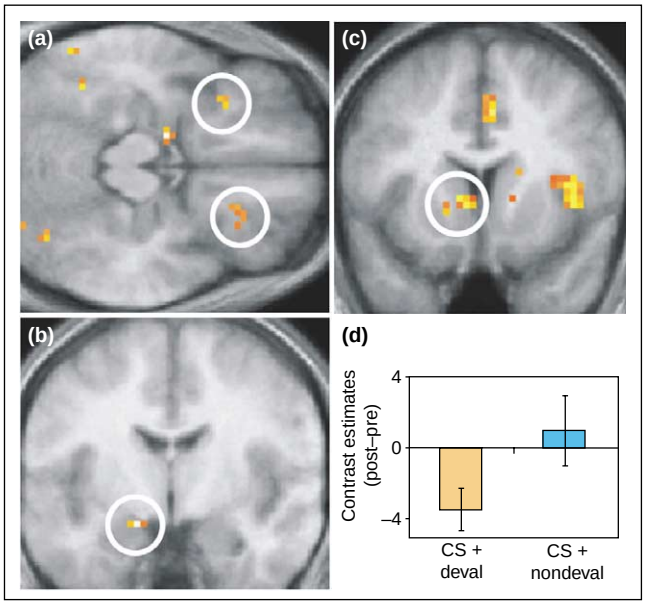
The role of the amygdala in coding affective value in general has recently been called into question. In two different studies, one in the olfactory and the other in the gustatory domain, stimuli matched for intensity (although differing in valence) were compared with those matched for valence (although differing in intensity) to distinguish areas involved in processing valence from those involved in processing intensity [4,7••]. Although responses in OFC were associated with valence, the amygdala was found to respond to intensity and not valence. These results could be taken to suggest a primarily sensory rather than affective role for the amygdala. However, such results appear to contradict evidence of an affective role for the amygdala from animal lesion and human neuropsychology studies [20••,21]. Furthermore, reward value can depend not only on the nature of the reward itself but also on the amount of that reward available, in the sense that more (in terms of concentration or quantity) of a given reward has a greater value than less of a given reward (although this is unlikely to be a linear association but dependent upon the utility to the animal of different amounts of that reward [22]). Thus, value is an interaction between valence and intensity rather than being synonymous with valence alone. In addition, future studies might need to forego subjective rating scale methodology, which are arguably indirect and rather insensitive measures of reward value, for more direct measures such as rate of instrumental responding or preference judgements (see [10,23•]).

Representation of predictive reward value

In addition to responding to rewarding and punishing stimuli once they have occurred, it is advantageous to be able to predict in advance when and where such rewards or punishments will occur so that behaviour can be organised prospectively. Fortunately, animals can make use of statistical regularities in their environment to form such predictions [24]. Value prediction in its simplest form can be studied by means of classical conditioning, which involves the presentation of an arbitrary neutral stimulus followed in a contingent fashion by a reward or punishment. After learning, the arbitrary stimulus takes on predictive value.

Neuroimaging studies have implicated amygdala, OFC and ventral striatum in reward prediction [25–27]. It is also important, however, to consider the content of predictive value representations in these brain areas. A conditioned stimulus can be associated with different aspects of an associated unconditioned stimulus, such as its sensory properties, its general affective properties (rewarding or aversive), or its specific reward value. To delineate predictive representations that access the specific value of the associated reward, a study was performed in which subjects were scanned whilst being presented with predictive cues associated with one of two food-

Figure 2



Predictive reward value coding in orbitofrontal cortex, amygdala and striatum. Results are shown from a classical conditioning paradigm in which arbitrary visual cues were paired with two food-related odours. Following devaluation of one of the odours (by feeding to satiety on the associated food), neural responses to the predictive cue associated with the devalued odour decreased selectively from pre to post-satiety **(a)** in orbitofrontal cortex, **(b)** amygdala and **(c)** ventral striatum. **(d)** An activity plot is shown from one of the regions (amygdala) illustrating the relative difference in activity (from pre to post satiety) for the cue associated with the devalued odour (CS + deval) and the cue associated with the non-devalued odour (CS + nondeval). This indicates that predictive value responses in these regions are linked to the specific value of the associated reward. (Adapted with permission from [28].)

related odours. Responses were compared to the cues before and after devaluation of one of the associated odours using selective satiation. Brain regions showing responses to predictive cues that tracked changes in reward value of the corresponding odours included OFC, amygdala and ventral striatum, indicating that predictive representations in those regions are linked to the specific value of the corresponding reward (Figure 2; [28]).

Distinct representations for prediction and receipt of reward

A related issue is whether predictive stimuli access the same or distinct neural representations as those elicited by the reward itself. This revisits a long-standing debate in animal learning concerning the nature of conditioned associations. According to stimulus-substitution theory, a conditioned stimulus (CS) acquires value by eliciting the same responses that would otherwise have occurred to the unconditioned stimulus (UCS) [29], in effect acting as a substitute for the unconditioned stimulus itself. Yet, it has also long been known that some conditioned responses are distinct from those elicited by the UCS suggesting that a CS is not merely a substitute for a UCS, but rather has its own unique properties [30]. Neuroimaging studies tend to support the notion of CS-unique representations in some brain areas. Specifically, ventral striatum and amygdala have been found to respond to predictors of reward and not to the reward itself after learning has taken place [25,26,31•]. However, such studies leave open the possibility that responses in these areas occur to the reward itself when unpredicted (before learning) but shift to the CS during the course of learning — a form of stimulus substitution. The critical test for this will be to compare activations elicited by a reward-predicting stimulus after learning to those elicited by the reward itself before learning.

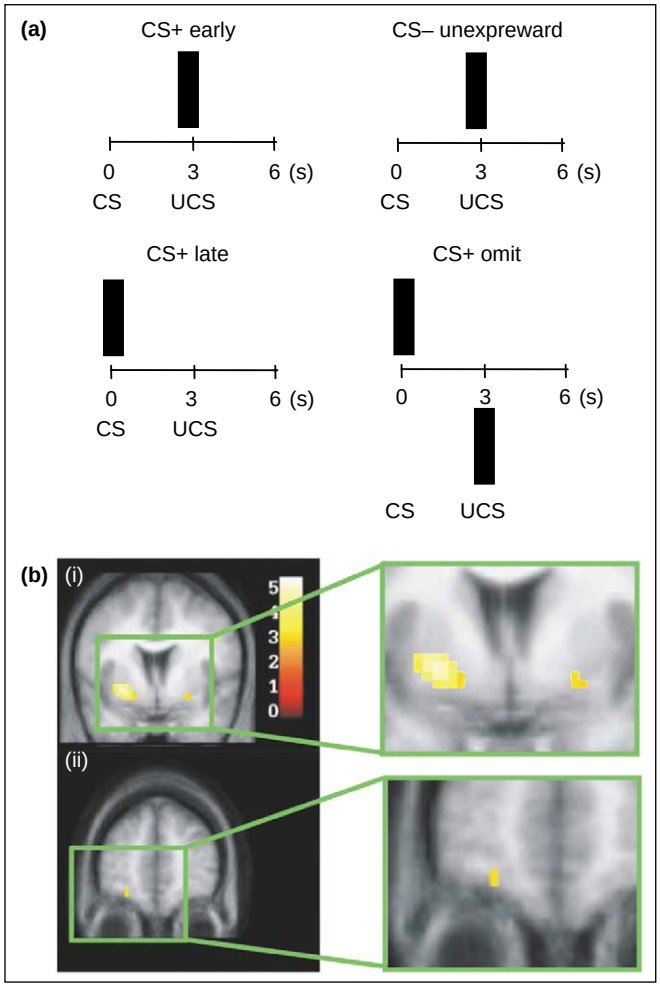
Computational mechanisms by which reward predictions are learned

How does the brain acquire predictive value representations? Some contemporary models of animal learning consider that learning occurs through a prediction error that signals discrepancies between expected and actual reward (or punishment) [32]. In one variant of this theory — temporal difference learning — predictions are formed about the expected future reward in a trial, and a prediction error reports differences in successive predictions of future reward [33]. Single unit studies in non-human primates implicate the phasic activity of dopamine neurons as a possible neural substrate of this signal [34]. The characteristics of this signal are reviewed elsewhere [35]. Briefly, over the course of learning the signal shifts its responses from the reward to the CS. Unexpected omission of reward results in a decrease in activity from baseline (a negative prediction error), whereas unexpected presentation of reward results in an increase in activity (positive prediction error). Human neuroimaging studies of classical conditioning for reward report prediction error signals in prominent target areas of dopamine neurons, namely ventral putamen and OFC (Figure 3; [36,37•]). These functional magnetic resonance imaging (fMRI) signal changes might reflect an interaction of intrinsic processing in those regions with the phasic activity of afferent dopamine neurons. Indeed, dopamine release has been reported in the striatum during reward prediction using positron emission tomography (PET) ligand measures [38•]. Dopamine neurons could facilitate learning of value predictions in these areas by gating plasticity between sensory and reward representations.

Salience versus reward in the striatum

Activation of the striatum has been reported during reward prediction, tracking reward prediction errors and in more complex gambling paradigms [25,36,39]. Recently it has been proposed that the striatum is involved in coding stimulus saliency rather than having an exclusive role in reward processing *per se* [40]. A similar

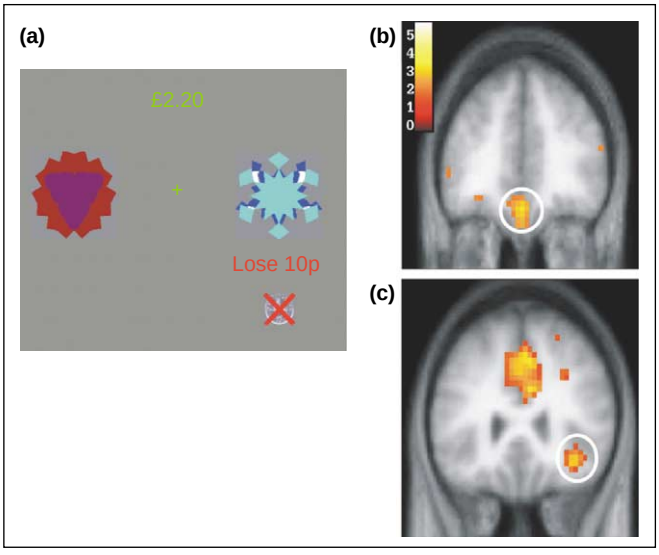
Figure 3



Prediction error signals in human striatum (ventral putamen) and orbitofrontal cortex during a classical conditioning paradigm in which in one trial type (CS+) an arbitrary visual cue is associated 3 s later with delivery of a taste reward (1M glucose), and in another trial type (CS−) a different cue is followed by no taste. In addition, occasional 'surprise' trials occur in which the CS+ is presented but the reward is omitted (CS and omit), and the CS− is presented but a reward is unexpectedly delivered (CS− unexpreward). **(a)** Schematic of putative temporal difference prediction error (PE) signals during the experiment. During early CS+ trials (before learning is established) the PE signal should occur at the time of delivery of the reward, whereas by late CS+ trials (post-learning) the signal should have switched to the time of presentation of the CS. On CS and omit trials a positive PE signal should occur at the time of presentation of the CS, but a negative PE signal should occur at the time the reward was expected (CS and omit). CS− and unexpreward trials should be associated with a positive signal at the time the reward is presented. **(b)** Parts of human ventral striatum (i) and orbitofrontal cortex (ii) showing a significant correlation with the temporal difference prediction error signal. (Data used with permission from [36].)

proposal has previously been made to account for dopamine function [41]. With regard to the striatum, this argument is supported by studies in which striatal activity is reported to 'non-rewarding' salient events such as presentation of infrequent distractor stimuli, as well as during an active reward task in which subjects must respond to obtain reward, compared with a passive task in which no response is required for reward [40,42]. A key issue in the reward specificity versus salience debate for dopamine neurons is whether or not such neurons also respond to equally salient punishing events. A similar question can be asked of the striatum. Indeed, there is strong evidence to implicate the ventral striatum in aversive as well as reward processing, especially during pain or anticipation of pain [43,44•]. These results at first sight support a saliency role for striatum. However, the omission of an unexpected reward produces deactivations in at least some parts of ventral striatum [25,36,37•,45]. An unexpected omission of reward is equally if not more salient than an unexpected delivery of reward. Yet these two types of salient event produce opposite response patterns in the striatum — a result not easily explicable in terms of a saliency hypothesis. Furthermore, a recent study has explored the role of this region in reporting

Figure 4



Decision making correlates in human OFC and adjacent ventral prefrontal cortex. This figure shows results from a study of visual discrimination reversal learning in which subjects have to choose between two stimuli, in which one stimulus is advantageous (associated with accumulating monetary reward) and the other is disadvantageous (associated with accumulating monetary loss). Subjects learn to choose the advantageous stimulus. However, after a period of time contingencies reverse and subjects must then switch their choice of stimulus. **(a)** Task illustration responses in some parts of OFC and adjacent anterior insula are related to subjects' behaviour on the subsequent trial. **(b)** Some areas respond if on the next trial subjects continue to choose the currently selected stimulus, **(c)** whereas other areas respond if on the subsequent trial subjects switch their choice of stimulus. (Adapted with permission from [54], Copyright 2003 by the Society for Neuroscience.)

temporal difference prediction errors for aversive learning (with pain) [46••]. As is the case with reward learning, unexpected omission of a cue associated with subsequent punishment produced a deactivation in the striatum. This again suggests that although this region is involved in both appetitive and aversive learning, it is not merely mediating stimulus salience. It should be noted that this does not necessarily imply an exclusive role for striatum in affective processing, nor does it rule out a role for striatum in saliency coding. It remains possible that both types of process co-exist within the structure, given the heterogeneous response profile of striatal neurons at the single neuron level [47].

Action selection for reward

The ability to form predictions of reward is only half the story. It is necessary to be able to act on those predictions. In a given context, specific actions might need to be performed to obtain reward. This requires learning of stimulus–response, or response–reward associations. Recent imaging studies have begun to explore brain mechanisms mediating this instrumental component. Such studies implicate the dorsal striatum, which shows activation when a contingency is established between responses and reward [48•,49,50] or even where there is merely a perceived contingency [51••]. These results are compatible with a role for the striatum (especially its dorsal aspects) in stimulus–response learning and suggest that in a manner analogous to stimulus-reward learning, such learning could be mediated by afferent dopamine input so that responses associated with greater predicted reward in a given context become reinforced and are thus more likely to be selected in future [52]. Action preparation for reward could also modulate activity in other brain regions such as lateral prefrontal and premotor cortex [53].

Decision making

To choose between different actions it is necessary to maintain a representation of the predicted future reward associated with each action. Such predictions then need to be compared and evaluated to select the action with the highest overall predicted reward value. This process is more complicated than at first sight, because estimations of predicted reward vary in their quality and depend on the number of samples of that action in the past as well as the variance of the reward distribution. This introduces dilemmas such as exploration versus exploitation — namely how long should be spent sampling different actions to gain a good estimate of predicted reward versus exploiting a particular action known to lead to a certain level of reward [22]. Neuroimaging studies have yet to breach these complexities, but simpler and more constrained decision making paradigms have been conducted [54,55]. One such paradigm is visual-discrimination reversal learning in which subjects have a choice of two stimuli, one of which if selected leads to accumulating monetary

gain whereas the other leads to accumulating loss. Subjects need to work out which is the advantageous stimulus and continue to choose it until contingencies reverse, after which they should switch their choice of stimulus. During performance of this task, some OFC regions respond if subjects choose the same stimulus on the next trial, whereas other regions respond if on the next trial subjects switch their choice of stimulus (Figure 4; [54]). Thus, a neural correlate of behavioural choice is present in OFC even in advance of that choice being implemented. Taken together with lesion data, these findings suggest a role for OFC in the decision making process itself [56].

Conclusions

Findings from neuroimaging studies indicate that brain regions such as OFC, amygdala and ventral striatum are involved in coding stimulus–reward value, maintaining representations of predicted future reward and future behavioural choice and might also play a part in integrating and evaluating reward predictions to guide decisions.

Future work will need to further differentiate the functions of each of these regions. For instance, one suggestion is that amygdala is involved in initial acquisition of predictions, but that OFC maintains more flexible representations that are updated following changes in contingencies [57,58]. Another view emphasises OFC contributions to guiding behaviour, in contrast with the amygdala, which is suggested to not be involved directly in behavioural choice [23•]. Furthermore, complex reward-related behaviours are not supported by any one of these areas in isolation, but are crucially dependent on interactions between these areas [20••,59]. Yet, such interactions have been neglected in the imaging literature to date. Recent developments in imaging methodology afford the opportunity to begin to characterize such interactions [60].

The focus in this review has been on a network of brain regions that have long been implicated in motivational processing. However, recent studies in non-human primates report reward-related neuronal responses in many other parts of the brain, including dorsolateral prefrontal cortex, anterior and posterior cingulate, and parietal cortex [61–64]. This suggests that reward-related information is present in many different brain regions, perhaps reflecting that implementation of complex goal-oriented behaviour requires recruitment of diverse cognitive resources. Thus, unravelling the neural mechanisms of reward might provide insight into fundamental principles of brain function.

Acknowledgements

Thanks to B Seymour and U Noppeney for their comments on this manuscript. Thanks also to R Dolan, P Dayan, J Gottfried and J Winston for helpful discussions.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Rolls ET, Sienkiewicz ZJ, Yaxley S: **Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey**. *Eur J Neurosci* 1989, **1**:53-60.

2. Critchley HD, Rolls ET: **Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex**. *J Neurophysiol* 1996, **75**:1673-1686.

3. O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F: **Representation of pleasant and aversive taste in the human brain**. *J Neurophysiol* 2001, **85**:1315-1321.

4. Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T: **Dissociation of neural representation of intensity and affective valuation in human gustation**. *Neuron* 2003, **39**:701-711.

5. Gottfried JA, Deichmann R, Winston JS, Dolan RJ: **Functional heterogeneity in human olfactory cortex: an event-related functional magnetic resonance imaging study**. *J Neurosci* 2002, **22**:10819-10828.

6. Rolls ET, Kringelbach ML, De Araujo IE: **Different representations of pleasant and unpleasant odours in the human brain**. *Eur J Neurosci* 2003, **18**:695-703.

7. •• Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Gabrieli JD, Sobel N: **Dissociated neural representations of intensity and valence in human olfaction**. *Nat Neurosci* 2003, **6**:196-202.
This fMRI study attempts to distinguish between brain areas processing valence and those processing intensity of odours, by matching some odours for intensity (although varying in valence) and others for valence (although varying in intensity). OFC is implicated in coding for valence, whereas the amygdala is proposed to be involved exclusively in processing odour intensity. This could suggest a primarily sensory rather than affective role for amygdala in stimulus processing. Interestingly, OFC responses in medial OFC were related to positive valence odours, whereas OFC responses in lateral OFC were related to negative valence odours.

8. Rolls ET, O'Doherty J, Kringelbach ML, Francis S, Bowtell R, McGlone F: **Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices**. *Cereb Cortex* 2003, **13**:308-317.

9. Blood AJ, Zatorre RJ, Bermudez P, Evans AC: **Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions**. *Nat Neurosci* 1999, **2**:382-387.

10. Aharon I, Etcoff N, Ariely D, Chabris CF, O'Connor E, Breiter HC: **Beautiful faces have variable reward value: fMRI and behavioral evidence**. *Neuron* 2001, **32**:537-551.

11. O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ: **Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness**. *Neuropsychologia* 2003, **41**:147-155.

12. Kawabata H, Zeki S: **Neural correlates of beauty**. *J Neurophysiol* 2004, **91**:1699-1705.

13. Elliott R, Newman JL, Longe OA, Deakin JF: **Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study**. *J Neurosci* 2003, **23**:303-307.

14. Rolls BJ, Rolls ET, Rowe EA, Sweeney K: **Sensory specific satiety in man**. *Physiol Behav* 1981, **27**:137-142.

15. O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F, Kobal G, Renner B, Ahne G: **Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex**. *Neuroreport* 2000, **11**:893-897.

16. Kringelbach ML, O'Doherty J, Rolls ET, Andrews C: **Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness**. *Cereb Cortex* 2003, **13**:1064-1071.

17. Morris JS, Frith CD, Perrett DI, Rowland D, Young AW, Calder AJ, Dolan RJ: **A differential neural response in the human amygdala to fearful and happy facial expressions**. *Nature* 1996, **383**:812-815.

18. Zald DH, Pardo JV: **Emotion, olfaction, and the human amygdala: amygdala activation during aversive olfactory stimulation**. *Proc Natl Acad Sci USA* 1997, **94**:4119-4124.

19. Canli T, Sivers H, Whitfield SL, Gotlib IH, Gabrieli JD: **Amygdala response to happy faces as a function of extraversion**. *Science* 2002, **296**:2191.

20. •• Holland PC, Gallagher M: **Amygdala-frontal interactions and reward expectancy**. *Curr Opin Neurobiol* 2004, **14**:148-155.
This excellent review covers similar ground to the present review, but discusses in more detail the role of amygdala and OFC in reward prediction, encompassing recent findings from the animal literature as well as human neuroimaging.

21. Adolphs R: **Neural systems for recognizing emotion**. *Curr Opin Neurobiol* 2002, **12**:169-177.

22. Dayan P, Abbott L: **Classical conditioning and reinforcement learning**. In *Theoretical Neuroscience*. Cambridge, MA: MIT Press; 2001:331-358.

23. • Arana FS, Parkinson JA, Hinton E, Holland AJ, Owen AM, Roberts AC: **Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection**. *J Neurosci* 2003, **23**:9632-9638.
In this PET study subjects were scanned while perusing fictional restaurant menus containing either high or low incentive items (based on individual preferences). In one condition subjects had to choose specific items, whereas in another condition they merely had to read the menu. OFC activity was related to choice of menu items (in some instances modulated by incentive value), whereas amygdala responses were related to incentive value, but independent of behavioural choice.

24. Gallistel CR: **Conditioning from an information processing perspective**. *Behav Processes* 2003, **62**:89-101.

25. Knutson B, Fong GW, Adams CM, Varner JL, Hommer D: **Dissociation of reward anticipation and outcome with event-related fMRI**. *Neuroreport* 2001, **12**:3683-3687.

26. O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ: **Neural responses during anticipation of a primary taste reward**. *Neuron* 2002, **33**:815-826.

27. Gottfried JA, O'Doherty J, Dolan RJ: **Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging**. *J Neurosci* 2002, **22**:10829-10837.

28. Gottfried JA, O'Doherty J, Dolan RJ: **Encoding predictive reward value in human amygdala and orbitofrontal cortex**. *Science* 2003, **301**:1104-1107.

29. Pavlov IP: *Conditioned reflexes*. Oxford: Oxford University Press; 1927.

30. Zener K: **The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response**. *Am J Psychol* 1937, **50**:384-403.

31. • Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D: **A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI**. *Neuroimage* 2003, **18**:263-272.
In this and previous studies from the same authors (e.g. [25]), a dissociation is proposed between regions involved in reward prediction and those in reward receipt. Whereas ventral striatum is implicated in reward prediction, medial prefrontal cortex is suggested to be involved in responding to receipt of the reward itself.

32. Rescorla RA, Wagner AR: **A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement**. In *Classical Conditioning II: Current Research and Theory.* Edited by Black AH, Prokasy WF. New York: Appleton Crofts; 1972:64-99.

33. Sutton RS, Barto AG: **Time derivative models of Pavlovian reinforcement**. In *Learning and Computational Neuroscience: Foundations of Adaptive Networks.* Edited by Gabriel M, Moore J. Cambridge, MA: MIT Press; 1990:497-537.

34. Schultz W, Dayan P, Montague PR: **A neural substrate of prediction and reward**. *Science* 1997, **275**:1593-1599.

35. Schultz W: **Predictive reward signal of dopamine neurons**. *J Neurophysiol* 1998, **80**:1-27.

36. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ: **Temporal difference models and reward-related learning in the human brain**. *Neuron* 2003, **38**:329-337.

37. • McClure SM, Berns GS, Montague PR: **Temporal prediction errors in a passive learning task activate human striatum**. *Neuron* 2003, **38**:339-346.
In this study prediction errors are induced not by unexpected omission or delivery of reward, but rather by varying when in the trial the reward is delivered. Delivery of reward later in a trial than expected produced activity in the putamen (albeit more dorsally than in O'Doherty *et al.* [36]). Sensitivity to stimulus timing is a key feature of temporal difference models of classical conditioning.

38. • Zald DH, Boileau I, El Dearedy W, Gunn R, McGlone F, Dichter GS, Dagher A: **Dopamine transmission in the human striatum during monetary reward tasks**. *J Neurosci* 2004, **24**:4105-4112.
The authors use raclopride PET to demonstrate dopamine release in the human striatum during performance of a reward-learning task.

39. Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA: **Tracking the hemodynamic responses to reward and punishment in the striatum**. *J Neurophysiol* 2000, **84**:3072-3077.

40. Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS: **Human striatal response to salient nonrewarding stimuli**. *J Neurosci* 2003, **23**:8092-8097.

41. Horvitz JC: **Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events**. *Neuroscience* 2000, **96**:651-656.

42. Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS: **Human striatal responses to monetary reward depend on saliency**. *Neuron* 2004, **42**:509-517.

43. Becerra L, Breiter HC, Wise R, Gonzalez RG, Borsook D: **Reward circuitry activation by noxious thermal stimuli**. *Neuron* 2001, **32**:927-946.

44. • Jensen J, McIntosh AR, Crawley AP, Mikulis DJ, Remington G, Kapur S: **Direct activation of the ventral striatum in anticipation of aversive stimuli**. *Neuron* 2003, **40**:1251-1257.
This study shows that ventral striatum is activated not only by predictors of reward but also by predictors of aversive stimuli, which implies that this region is not reward specific but has a more general role in motivational processing.

45. Pagnoni G, Zink CF, Montague PR, Berns GS: **Activity in human ventral striatum locked to errors of reward prediction**. *Nat Neurosci* 2002, **5**:97-98.

46. •• Seymour B, O'Doherty JP, Dayan P, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS: **Temporal difference models describe higher-order learning for pain**. *Nature* 2004, **429**:664-667.
The authors demonstrate that temporal difference prediction error signals are present in ventral striatum not only during reward prediction but also during prediction of aversive stimuli (pain). Furthermore, prediction errors are shown following unexpected delivery or omission of learned cues rather than following delivery or omission of the pain itself.

47. Schultz W, Tremblay L, Hollerman JR: **Changes in behavior-related neuronal activity in the striatum during learning**. *Trends Neurosci* 2003, **26**:321-328.

48. • Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M: **A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task**. *J Neurosci* 2004, **24**:1660-1665.
This study shows that during action selection for reward, activity in dorsal striatum correlates with a measure of learning as assessed by the degree of convergence between subjects' responses and optimal behaviour in the task.

49. Elliott R, Newman JL, Longe OA, William Deakin JF: **Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems**. *Neuroimage* 2004, **21**:984-990.

50. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ: **Dissociable roles of ventral and dorsal striatum in instrumental conditioning**. *Science* 2004, **304**:452-454.

51. •• Tricomi EM, Delgado MR, Fiez JA: **Modulation of caudate activity by action contingency**. *Neuron* 2004, **41**:281-292.
The authors show that dorsal striatum activity occurs during action selection even if there is just a perceived contingency between action selection and reward (in actuality rewards were delivered non-contingently).

52. Montague PR, Dayan P, Sejnowski TJ: **A framework for mesencephalic dopamine systems based on predictive Hebbian learning**. *J Neurosci* 1996, **16**:1936-1947.

53. Ramnani N, Miall RC: **Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation**. *Cereb Cortex* 2003, **13**:318-327.

54. O'Doherty J, Critchley H, Deichmann R, Dolan RJ: **Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices**. *J Neurosci* 2003, **23**:7931-7939.

55. Rogers RD, Ramnani N, Mackay C, Wilson JL, Jezzard P, Carter CS, Smith SM: **Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition**. *Biol Psychiatry* 2004, **55**:594-602.

56. Bechara A, Tranel D, Damasio H: **Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions**. *Brain* 2000, **123**:2189-2202.

57. Sanghera MK, Rolls ET, Roper-Hall A: **Visual responses of neurons in the dorsolateral amygdala of the alert monkey**. *Exp Neurol* 1979, **63**:610-626.

58. Morris JS, Dolan RJ: **Dissociable amygdala and orbitofrontal responses during reversal fear conditioning**. *Neuroimage* 2004, **22**:372-380.

59. Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA: **Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex**. *J Neurosci* 2000, **20**:4311-4319.

60. Friston KJ, Harrison L, Penny W: **Dynamic causal modelling**. *Neuroimage* 2003, **19**:1273-1302.

61. Watanabe M, Hikosaka K, Sakagami M, Shirakawa S: **Coding and monitoring of motivational context in the primate prefrontal cortex**. *J Neurosci* 2002, **22**:2391-2400.

62. Tanji J, Shima K, Matsuzaka Y: **Reward-based planning of motor selection in the rostral cingulate motor area**. *Adv Exp Med Biol* 2002, **508**:417-423.

63. McCoy AN, Crowley JC, Haghighian G, Dean HL, Platt ML: **Saccade reward signals in posterior cingulate cortex**. *Neuron* 2003, **40**:1031-1040.

64. Platt ML, Glimcher PW: **Neural correlates of decision variables in parietal cortex**. *Nature* 1999, **400**:233-238.