Introduction

A reward is anything an animal or human will work to attain. Some rewards are primary reinforcers in that they have innate (or unlearned) value to the organism, whereas other rewarding stimuli can be considered secondary reinforcers in that they have acquired value through being associated with primary reinforcers. Seeking out rewards (or avoiding punishers) forms the basis of most goal-directed behaviors. Therefore, understanding how rewards are processed in the brain is central to understanding the neural basis of behavior. In the last 15 years, modern neuroimaging techniques in humans such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have complemented single-unit neurophysiology and lesion study approaches in nonhuman animals as a means of providing insight into the neural mechanisms underlying reward processing in the human brain. In this article, the contribution of human neuroimaging studies to our current understanding of neural processes underlying reward is reviewed.

Key Brain Regions Involved in Reward Processing

Although reward-related modulation of neural activity has been found to be widespread in the brain, having been reported in the parietal cortex, dorsal prefrontal cortex, and premotor and posterior cingulate cortices, among other areas, in this article we focus on the functions of a core network of brain areas that play a key role in reward processing. These areas are the amygdala in the medial temporal lobes, the orbitofrontal cortex on the ventral surface of the frontal lobes, the adjacent medial prefrontal cortex, the ventral and dorsal striatum, and the dopaminergic neuron-containing nuclei in the midbrain.

Technical Challenges Facing Human Imaging Studies of Reward Processing

Studying reward presents a relatively unique set of challenges to the imaging neuroscientist. This arises due to an interaction between the physiological properties of the reward system and the properties of the imaging techniques themselves. First, the vast majority of functional neuroimaging studies at present make use of the blood oxygenation level-dependent contrast (BOLD) in magnetic resonance imaging (MRI). This measure depends on detecting subtle differences in local magnetic susceptibility effects caused by neural activity-induced changes in the concentration of deoxy-hemoglobin, as opposed to oxyhemoglobin. Yet many of the brain areas that are the natural focus of reward studies – namely the orbitofrontal cortex, amygdala, and ventral striatum – happen to be at the border of areas with large susceptibility differences between brain tissue and air in the sinuses. The presence of these large extrinsic susceptibility differences can produce dropout or loss of signal as well as spatial distortion in the BOLD contrast images. A number of techniques have been developed to deal with this problem which can recover much, although not all, of the signal loss in these areas; these include tilted acquisition (typically at 30° to the anterior commissure–posterior commissure (AC–PC) line), application of a compensation gradient prepulse, and the use of a short echo time. An alternative approach is to use PET, which does not suffer from dropout, but the poor temporal and spatial resolution of that technique substantially limits its usefulness. Other MRI techniques such as perfusion MRI may also be less affected by dropout, although at the expense of sensitivity. A second challenge for the imaging of reward systems is that many aspects of reward processing may relate to the functions of specific neurotransmitter systems, including the dopamine system and opioid system, the specific effects of which cannot be isolated with standard hemodynamic imaging approaches. To study this in humans requires administration of specific targeted radioligand molecules which bind with the relevant receptors and that can then be recorded using PET, for example, using the radioligand raclopride to image the dopamine system. Alternatively, the effects of pharmacological modulation of these neurotransmitter systems (by the systemic administration of an agonist or antagonist) on neural activity can be assayed with standard hemodynamic imaging techniques.

Distinct Components of Reward Processing

Reward processing can usefully be fractionated into four distinct components: (1) neural responses pertaining to receipt of a rewarding outcome, (2) neural responses related to expectation or anticipation of a rewarding outcome, (3) neural signals underlying
learning of reward expectancies, and (4) neural responses related to action choice for reward. Here we discuss the role of the different brain regions in the reward network in mediating these distinct components.

**Receipt of a Rewarding Outcome**

In order to survive, an animal needs to be able to quickly evaluate the affective significance of stimuli in its environment to determine whether such stimuli are beneficial (rewarding) or harmful (punishing) so that appropriate approach or avoidance behavior can be initiated. This raises perhaps the most fundamental question about how rewards might be processed in the brain: Where in the sensory processing pathway is the reward (or punishment) value of a stimulus encoded? Rewarding or punishing stimuli can be present in each of the five sensory modalities, each of which have dedicated sensory-processing pathways. One possibility is that reward value is encoded within the sensory pathway of each modality separately, such that there are separate neural apparatuses for decoding reward value in each sensory pathway. Alternatively, the decoding of reward value may occur in multimodal brain area(s) such that a single dedicated apparatus accomplishes this for each sensory modality. Substantial evidence has now accrued from single-unit neurophysiology studies to support the latter hypothesis, which is that one or two brain areas appear to be involved in encoding the reward value of stimuli from each sensory modality. These areas are the orbitofrontal cortex and the amygdala. Both of these areas receive direct inputs from each of the sensory pathways and are therefore highly multimodal brain regions.

**Role of orbitofrontal cortex in encoding stimulus reward value**

Single-unit neurophysiology studies have reported evidence that neural responses in the orbitofrontal cortex to a stimulus input can be modulated as a function of changes in the value of that stimulus to the animal. Brain-imaging studies in humans also support the conclusion that there is a key role for this area in encoding stimulus reward value for a variety of rewards in different modalities. In a typical study, subjects are scanned while being presented with a range of different stimuli in a specific modality that are either pleasant, affectively neutral, or aversive to the subject. For instance, to study areas encoding stimulus reward value in the olfactory modality, subjects might be presented with a range of odors, some of which are perceived to be pleasant by the subject, some of which are neutral, and some of which are considered aversive. Direct comparisons can then be computed among areas activated in response to the pleasant, aversive, and neutral odors, such that areas involved in encoding positive reward value for an odor could be revealed by testing for areas showing increased activity to the pleasant compared to the neutral or aversive odors. Such studies have revealed increased activity in parts of the orbitofrontal cortex to the receipt of pleasant as opposed to neutral or aversive odors. Other regions of the orbitofrontal cortex have been identified as showing increased responses to aversive odors. A similar methodology has been used in studies exploring the neural encoding of value in other modalities, such as touch, taste, audition, and vision. An example of rewarding stimuli in the visual modality is attractive faces, which human subjects are prepared to work for in order to view, particularly those of the opposite sex. As in the olfactory domain, attractive faces produce significant activity in the orbitofrontal cortex (Figure 1(a)). This is also the case for rewards in each of the other sensory modalities that have been tested. An alternative approach to comparing neural responses to stimulus categories is to correlate individual subjective ratings of pleasantness for a given stimulus with neural activity in response to that stimulus to determine the regions that correlate positively with perceived pleasantness. In the case of the attractive face example, this approach again reveals strong activity in the orbitofrontal cortex that correlates positively with subjective ratings of pleasantness for a stimulus.

Yet another approach is to make use of the phenomenon of selective satiation or sensory-specific satiety. This is the phenomenon whereby if a particular food is eaten to satiety, the pleasantness of that food can selectively decrease while the pleasantness of other foods not eaten can remain pleasant. An example of this is a person eating a main course of spaghetti in a restaurant until satiated but still having room for an additional dessert course of, say, tiramisu. The spaghetti has gone from being initially pleasant to having little or no reward value (and may even be aversive if too much is consumed), whereas the tiramisu still has high reward value until it too is consumed to satiety. Neural encoding of reward value can be studied in this way by presenting subjects with two or more food stimuli considered pleasant by a hungry subject, measuring neural activity elicited by these stimuli, and then feeding the subject to satiety on only one of the foods. If, following this feeding procedure, the subject is again scanned while being presented with the two foods, the food eaten to satiety will no longer be pleasant, whereas the food not eaten will remain pleasant. By testing for brain regions responding to the food stimuli that decrease selectively to the food eaten postsatiety, it is possible to identify regions encoding reward value. Once again, these studies reveal
activity in the orbitofrontal cortex tracking changes in the perceived reward value of stimuli (for both olfactory food rewards and for whole-food stimuli).

Some rewards are not clearly tied to specific modalities, but are more abstract in nature. A paradigmatic example of this is monetary reward, the receipt of which can be conveyed in multiple modalities. Consistent with other types of rewarding stimuli, monetary rewards also activate the orbitofrontal cortex (Figure 1(b)). Furthermore, the magnitude of activation in this area scales with the magnitude of the points received (right). fMRI, functional magnetic resonance imaging. (a) Adapted from O’Doherty J, Winston J, Critchley H, Perrett D, Burt DM, and Dolan RJ (2006) Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. Neuropsychologia 41(2): 147–155. (b) Adapted from Daw ND, O’Doherty JP, Dayan P, Seymour B, and Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. Nature 441(7095): 876–879.

Figure 1 Representation of stimulus–reward value in human orbitofrontal cortex. (a) fMRI study in which subjects were presented with faces that were high and low in attractiveness; (b) fMRI study in which subjects made choices in order to win arbitrary points which they were informed would be converted into money after the experiment. In (a), attractive faces are a class of rewarding visual stimuli that activate the orbitofrontal cortex (example faces from the high-attractive category are shown on the left). A region of medial orbitofrontal cortex showed increased activity during presentation of high-compared to low-attractive faces (right and middle). In (b), the orbitofrontal cortex also activates in response to abstract rewards such as money or arbitrary points. The medial orbitofrontal cortex was found to respond at the time of reward receipt (left and middle). Moreover, the magnitude of activity in this region scaled according to the magnitude of the points received (right). fMRI, functional magnetic resonance imaging. (a) Adapted from O’Doherty J, Winston J, Critchley H, Perrett D, Burt DM, and Dolan RJ (2006) Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. Neuropsychologia 41(2): 147–155. (b) Adapted from Daw ND, O’Doherty JP, Dayan P, Seymour B, and Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. Nature 441(7095): 876–879.

Amygdala and reward In the early neuroimaging literature, the amygdala was often implicated in aversive learning such as fear conditioning, in the perception of fearful or angry faces, or in responding to aversive tastes or odors. This led to an initial emphasis on the role of the human amygdala in negative affect. However, more recently the amygdala has also been implicated in responding to rewarding stimuli in different modalities, including pleasant tastes, odors, and monetary gain, suggesting that, consistent with the results of single-unit neurophysiology and lesion studies in animals, the amygdala is involved in appetitive as well as aversive processing.

However, the role of the amygdala in encoding either positive or negative valence, has recently been challenged. In two separate studies, one involving gustatory reward and the other olfactory, the intensity and valence of a stimulus were varied separately such that some stimuli were rated by subjects as intense but not pleasant, whereas other stimuli were rated as pleasant but not intense. These studies found that the amygdala responded to the intensity but not the pleasantness of the stimuli. However, a further investigation showed that the amygdala cares not only about intensity but also about the interaction between intensity and valence, such that it responds
more to a highly intense but highly pleasant stimulus than it does to a less intense but still highly pleasant stimulus. This suggests that the amygdala is involved in encoding stimulus reward value, which is the combination of intensity (or magnitude) and valence rather than either valence or intensity alone.

**Expectation of Reward**

The ability to predict when and where rewarding or punishing stimuli will arrive confers considerable advantage to an animal because it enables appropriate behavior to be organized prospectively. For this reason, humans and other animals have developed sophisticated neural mechanisms to exploit statistical structure in the environment to learn to predict the arrival of such stimuli. Classical or Pavlovian conditioning is the simplest paradigm for studying reward prediction in the laboratory. In Pavlovian conditioning, an arbitrary neutral stimulus takes on predictive value through being repeatedly paired with the subsequent delivery of a behaviorally significant stimulus. Human imaging studies have used Pavlovian conditioning procedures to assay brain regions involved in encoding predictions of future reward. In a typical study, subjects are scanned while being presented with different visual cues (such as fractal images). One of these cues (designated the CS+) is paired repeatedly with subsequent delivery of reward (such as a pleasant odor or a pleasant taste or juice stimulus), whereas another cue (the CS−) is either paired with nothing or with an affectively neutral stimulus. Neural responses related to reward prediction can then be isolated by comparing activity to the onset of the CS+ compared to the CS−. Such studies have revealed significant responses related to reward prediction in both the orbitofrontal cortex and amygdala (Figure 2), suggesting that these regions are involved not only in responding to the receipt of a reward itself but also in anticipating the subsequent delivery of reward. Another region found to show strong expected reward responses is the ventral striatum, composed of the nucleus accumbens and adjacent ventral putamen. Expected reward signals have been found in each of these areas during anticipation of olfactory, taste, flavor, and monetary rewards.

**Neuroeconomic variables and expected reward** As already demonstrated, the application of paradigms borrowed from the field of behavioral psychology, such as Pavlovian conditioning, has proved useful in neuroscientific investigations of reward processing. Another fruitful avenue has been to bring to bear well-defined concepts from the field of economics as part of the emerging interdisciplinary research agenda known as neuroeconomics. Relevant concepts from economics include expected value, which formally is the product of the probability of obtaining a reward times the magnitude of the reward, and utility, which captures an individual’s revealed preferences as a function of the magnitude (quantity) or type of reward. Utility can be expressed as a function of reward magnitude and can be a nonlinear function thereof. For example, the marginal (or relative difference in) utility for an individual of obtaining $102 as opposed to $100 may be much less than the marginal utility of obtaining $4 as opposed to $2, even though the difference in objective reward magnitude is the same. When the preferences for that individual are plotted as a function of magnitude, a concave utility curve is obtained showing that there is a much shallower increase in utility at the $100–102 end of the scale than at the $2–4 end. Expected utility is simply the probability of reward times its utility. This concept can be used to explain differences in preference between risky and nonrisky gambles even if the expected values for the gambles are identical. Variations of utility and expected utility theory have been proposed to account for various anomalies in human preference behavior, such as prospect theory, which accounts for asymmetries in subjects’ propensity for risky decision making for gambles involving losses compared to gains.

These concepts have been applied in human neuroimaging studies to investigate neural responses related to expected value, expected magnitude, and expected probability. A number of studies have manipulated these variables by presenting visual cues that independently signal differing probabilities and magnitudes of subsequent reward. Correlations can then be computed between neural activity in response to the different cues and these variables to reveal areas responding separately to probability, magnitude, or the product of the two. Such studies revealed responses to expected magnitude, probability, and expected value in the ventral striatum and medial prefrontal cortex (Figure 3). As yet, no clear evidence has emerged of expected utility signals in the human brain, possibly because neural representations of expected utility may be difficult to dissociate experimentally from expected value because the signals may under many circumstances differ only subtly.

**Risk** Another variable relevant to anticipated reward is the variance in the distribution of expected reward, referred to as risk or uncertainty. Risk is a quadratic function of the probability of reward, such that risk is maximal at probability 0.5, when the delivery of reward is maximally uncertain, and minimal at 1.0 or 0 probability, when the reward is
either certain to occur or certain not to occur. Humans are known to vary in their sensitivity to risk; some individuals are risk averse, preferring a nonrisky gamble to a risky gamble if the gambles are equal in expected value, whereas other individuals are risk prone, favoring a risky over a nonrisky gamble with equal expected values. Other individuals are risk neutral in that they show no preference either way. Several fMRI studies have revealed increased activity in the anterior insula or lateral orbitofrontal cortex when subjects are presented with risky as opposed to nonrisky gambles. Given that these areas are often implicated in aversive processing, such as responding to pain or its anticipation, and to receipt of monetary loss, responses to risk in these areas could reflect the negative affective component of risk, at least in those subjects who are risk averse. Risk signals have also been found to be present in the tonic ramping activity of dopamine neurons in singleunit neurophysiology studies in nonhuman primates. Similar signals have been found in fMRI studies in target areas of dopamine neurons such as the ventral striatum; tonic ramping anticipatory activity in this area has been found to correlate

Figure 2  Neural responses during anticipation of reward: (a) neural responses during anticipation of a pleasant taste reward; (b) areas of the orbitofrontal cortex and amygdala showing anticipatory responses that track changes in the value of an associated olfactory reward induced by selective satiation. In (a), human subjects were presented with a visual cue that indicated subsequent delivery of either a pleasant sweet taste, a mildly unpleasant taste, or an affectively neutral control solution. Shown are the areas responding more to anticipation of the pleasant (left) compared to the mildly unpleasant taste (right). In (b), hungry subjects were conditioned to two cues which signaled the subsequent delivery of two food-related odors rewards (vanilla and peanut butter). Subjects were then fed to satiety on one of the foods, selectively devaluing one of the food-related odors. Subjects were then scanned again while being presented with the same two conditioned stimuli. Shown are the amygdala (top left) and orbitofrontal cortex (bottom left) responses to the cues that track changes in the value of the associated odors, indicating that these areas encode anticipatory reward value. The right plots show the difference in activity between pre- and postsatiety to conditioned cues associated with the odor of the food fed to satiety (Tgt CS + u) compared to response to cues associated with odor of food not fed to satiety (nTgt CS + u). (a) Adapted from O’Doherty J, Deichmann R, Critchley HD, and Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. Neuron 33(5): 815–826. (b) Adapted from Gottfried JA, O’Doherty JP, and Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301(5636): 1104–1107.
with risk or uncertainty (Figure 4). The presence of risk signals in addition to expected value could be important for behavioral decision making because, as discussed earlier, most individuals are sensitive not only to the expected value of a given choice but also to the variance of the expected reward outcome when making decisions.

**Learning Expected Reward Signals**

The finding of expected value signals in the brain raises the question of how such signals are learned in the first place. An influential theory from behavioral psychology put forward by Rescorla and Wagner suggests that learning of reward expectancy is mediated by the degree of surprise engendered when an outcome is presented or, more precisely, the difference between what is expected and what is received. Formally this is called a prediction error, which in the Rescorla–Wagner formulation can take on either positive or negative sign depending on whether an outcome is greater than expected (a positive prediction error signal) or less than expected (a negative prediction error signal). This prediction error is then used to update the expected reward associated with a particular stimulus or cue in the environment, so that, if this cue always precedes...
reward (and hence is fully predictive of reward), eventually the expected value of the cue will converge to the value of the reward, at which point the prediction error is zero and no further learning will take place. The role of prediction error signals in reward learning in the brain was first uncovered through observation of the phasic activity of dopamine neurons in nonhuman primates during the performance of simple conditioning tasks. These neurons were found to demonstrate a response profile approximately captured by real-time extensions of the Rescorla–Wagner model such as temporal difference learning. fMRI studies in humans have found evidence for prediction error signals in target regions of dopamine neurons such as the ventral striatum and orbitofrontal cortex during reward learning (Figure 5). A recent study attempted to directly link prediction error signals observed in fMRI to dopamine release by administering systemic dopamine agonists and antagonists to human subjects while they performed a simple reward-learning task. Prediction error signals in the striatum were found to be significantly affected by the pharmacological manipulation, suggesting that the generation of such signals in humans may be directly related to the activity of dopamine neurons.

**Action Selection for Reward**

Once predictions of future reward are established, the next step is to use these predictions to guide behavior. More specifically, humans and other animals need to be able to bias their action selection so that they choose those actions which in a given context lead to the greatest probability of obtaining future reward or avoiding punishers. This phenomenon is often referred to as instrumental conditioning and involves learning of stimulus–response and/or stimulus–response–outcome associations. Human neuroimaging studies of reward-based action selection have implicated the dorsal part of the striatum, which is activated when a contingency is established between responses and reward or even when such a contingency is merely perceived. Other regions that may play a role in reward-based action selection include the anterior cingulate cortex, premotor cortex, and lateral prefrontal cortex.

A major focus of current research is to gain an understanding of how the brain can make choices or decisions between different available actions. Human imaging studies have begun to probe the brain regions involved in implementing behavioral decisions using simple choice paradigms. Perhaps the simplest paradigm for this is the n-armed bandit task, in which a subject is presented with a number of different options, or bandits, each of which pays out reward with differing probabilities or magnitudes. One of the bandits is generally better than the others in that the consistent selection of that bandit will lead to more cumulative reward than is available from the others. The subjects’ objective is to establish which bandit

![Figure 5](https://example.com/figure5.png)

**Figure 5** Prediction error (PE) signals in the human striatum (ventral putamen) and orbitofrontal cortex: (a) schematic of putative temporal difference PE signals during the experiment; (b) parts of the human ventral striatum (top) and orbitofrontal cortex (bottom) that showed a significant correlation with the temporal difference PE signal. The study was a classical conditioning paradigm in which in one trial type (CS+) an arbitrary visual cue was associated 3 s later with delivery of a taste reward (1M glucose) and in another trial type (CS−) a different cue was followed by no taste. In addition, occasional surprise trials occurred in which the CS+ was presented but the reward was omitted (CS+ omit) and in which the CS− was presented but a reward was unexpectedly delivered (CS− unexpreward). In (a) schematics of putative temporal difference PE signals during the experiment are shown. During early CS+ trials (before learning was established), the PE signal should occur at the time of delivery of the reward, whereas by the late CS+ trials (postlearning) the signal should switch to the time of presentation of the CS. During CS+ omit trials, a positive PE signal should occur at the time of delivery of the reward, whereas by the late CS+ trials (postlearning) the signal should switch to the time of presentation of the CS+. During CS+ unexpreward 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− unexpreward trials should be associated with a positive signal at the time the reward was presented. CS, conditioned response; UCS, unconditioned response. Adapted from O’Doherty JP, Dayan P, Friston K, Critchley H, and Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. Neuron 38(2): 329–337, with permission from Elsevier.
pays the most and choose that consistently in order to maximize reward. However, the subject is not informed at the beginning which bandit is better but, instead, must work this out from trial and error by making selections in order to gain experience of the rewards available from the different bandits.

Useful insights into how humans or other animals might solve this type of problem have come from a branch of computer science known as reinforcement learning, which is concerned with building artificial algorithms for solving stochastic decision problems. The basic idea behind most reinforcement learning models is that in order to choose optimally between different actions, an agent needs to maintain internal representations of the expected reward available on each action and then subsequently choose the action with the highest expected value. Also central to these algorithms is the notion of a prediction error signal which is used to learn and update the expected values for each action through experience, just as in the Rescorla–Wagner learning model for Pavlovian conditioning described earlier. Consistent with the reinforcement learning framework, human imaging studies have found evidence for expected value signals in brain regions such as the medial and orbital prefrontal cortex and amygdala during the performance of simple variants of the $n$-armed bandit task (Figure 6). Moreover, prediction error signals have also been found in both the ventral and dorsal striatum and the prefrontal cortex during the performance of such tasks, consistent with the notion that such signals may be used to update action values as well as Pavlovian stimulus–outcome associations. These studies suggest that reinforcement learning

![Figure 6](image-url) Expected value signals in the orbitofrontal cortex, medial prefrontal cortex, and amygdala during decision making: (a) illustration of payoffs available after the choice of the correct or incorrect stimulus; (b) areas of the medial orbitofrontal cortex, medial prefrontal cortex, and amygdala correlating with expected reward value (derived from a simple computational model) at the time of choice; (c) fMRI responses in medial prefrontal cortex plotted against expected value to illustrate the strong colinearity between fMRI signals in this area and expected value. The task is a simple variant of a two-armed bandit task called probabilistic reversal learning in which a subject chooses between two stimuli which pay out monetary gains and losses on a probabilistic basis. One stimulus is correct, in that the continual choice of that stimulus leads to an overall monetary gain, whereas the other stimulus is incorrect, in that the choice of that stimulus leads to an overall monetary loss. After a time, the contingencies reverse so that the advantageous stimulus becomes the disadvantageous and vice versa. Subjects have to keep track of which stimulus is correct and switch their responses once they deem the contingencies have reversed. In (a), a monetary gain of 25 cents is signaled by the presentation of a US quarter dollar coin, whereas the monetary loss of 25 cents is signaled by the presentation of a red cross superimposed over a quarter dollar. fMRI, functional magnetic resonance imaging. Adapted from Hampton AN, Bossaerts P, and O’Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *Journal of Neuroscience* 26(32): 8360–8367, Copyright 2006 by the Society for Neuroscience.
models may provide a useful analogy to at least some of the neural processes underlying decision making in humans.

Perhaps unsurprisingly, such models are unlikely to account for all aspects of human choice behavior. For instance, these models are unable to learn about higher-order decision rules, such as interdependencies between actions and other exogenous variables including time, whereas expected value representations in the medial prefrontal cortex have been found to encode these decision rules. Furthermore, existing reinforcement learning models may not adequately account for how humans solve the exploration/exploitation dilemma, in which a decision maker needs to balance the need to choose what is currently believed to be the best action with the need to sample other available actions just in case those actions turn out to be more lucrative. A recent study showed that regions of the anterior prefrontal cortex may become engaged when subjects switch between exploratory and exploitative action selection modes, but such a putative switching mechanism is not easily accounted for in current reinforcement learning approaches to this problem. Finally, simple reinforcement learning models may not adequately account for the behavioral distinctions between goal-directed action choice, in which actions are selected with reference to the current incentive value of the associated outcome (involving response–outcome associations), and habit learning, whereby an action is selected purely on the basis of a learned stimulus–response association without reference to the outcome which that response engenders.

See also: Appetitive Systems: Amygdala and Striatum; Cognition: An Overview of Neuroimaging Techniques; Conditioning: Theories; Connectivity of Primate Reward Centers; Cortical Processing of the Reward Value of Food; Delayed Reinforcement: Economics; Delayed Reinforcement: Neuroscience; Neuroeconomics: History; Neuroimaging; Procedural Learning: Classical Conditioning; Reinforcement Models: Reward Decision-Making; Reward Neurophysiology and Orbifrontal Cortex; Reward Neurophysiology and Primate Cerebral Cortex; Reward Systems: Human.

Further Reading