



ELSEVIER

Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology

Wolfram Schultz

Neurons in a small number of brain structures detect rewards and reward-predicting stimuli and are active during the expectation of predictable food and liquid rewards. These neurons code the reward information according to basic terms of various behavioural theories that seek to explain reward-directed learning, approach behaviour and decision-making. The involved brain structures include groups of dopamine neurons, the striatum including the nucleus accumbens, the orbitofrontal cortex and the amygdala. The reward information is fed to brain structures involved in decision-making and organisation of behaviour, such as the dorsolateral prefrontal cortex and possibly the parietal cortex. The neural coding of basic reward terms derived from formal theories puts the neurophysiological investigation of reward mechanisms on firm conceptual grounds and provides neural correlates for the function of rewards in learning, approach behaviour and decision-making.

Addresses

Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK
e-mail: ws234@cam.ac.uk

Current Opinion in Neurobiology 2004, **14**:139–147

This review comes from a themed issue on
Cognitive neuroscience
Edited by John Gabrieli and Elisabeth A Murray

0959-4388/\$ – see front matter
© 2004 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.conb.2004.03.017

Introduction

Although rewards are environmental objects that interact with the brain through primary sensory receptors, they come in different sensory modalities and act on the brain without specialised receptors. Thus, the rewarding aspects of environmental stimuli need to be extracted from a variety of primary sensory sources. Rewards cannot be described by their physical properties alone. That is, the basic variables of reward information cannot be gleaned from the common parameters of sensory systems, such as size and position of receptive fields and intensity and contrast of stimuli. Rather, their function is defined by the behaviour of the organism and formalised by behavioural learning and motivational theories. Much progress has been made in understanding the basic parameters that determine the influence of rewards on the behaviour of the organism through the study of the underlying brain mechanisms. The present review is an attempt to summarise our knowledge of the basic com-

ponents of reward information that are extracted from multisensory stimuli and processed by the reward mechanisms of the brain. Despite the large variety of non-nutrient objects that constitute rewards for humans, this review primarily considers food and liquid rewards that can be delivered in a quantifiable way to monkeys, which are the main subjects for the studies used here.

Basic theoretical reward terms

According to animal learning theory, rewards have three basic functions in behaviour. First, they induce learning, as they make subjects come back for more (positive reinforcement); second, they induce approach and consummatory behaviour for acquiring the reward object; and third, they induce positive emotions, which are inherently difficult to investigate in animals. Rewards can serve as goals of behaviour if the reward and the contingency between action and reward are represented in the brain during the action [1]. By contrast, punishers induce avoidance learning, withdrawal behaviour and negative emotions.

As the intensity of behavioural reactions to rewards is graded, rewards appear to have motivational value that permits the organism to compare and choose among different rewards. The influence of rewards on behaviour depends in many instances on the vegetative drive state of the organism. However, additional factors determine reward value. For example, a subject performs a reward-directed action with certain intensity but reduces that intensity after being exposed to a more desirable object (negative contrast effect) [2], and this effect seems to be exacerbated when another animal receives visibly more reward for the same effort [3].

Reward-directed learning can occur by associating a stimulus with a reward (Pavlovian conditioning) or by associating an action with a reward (operant conditioning), with more complex associations among stimuli, actions and outcomes occurring during various stages of learning. Environmental stimuli and objects acquire reward value through past experience in a Pavlovian manner and therefore come to predict the associated rewards. It is usually difficult to distinguish exactly between the reward-predicting stimulus and the actual primary reward [4], and for simplicity we can call a reward an object that acts on the body through proximal somatosensory receptors, such as a piece of food or a drop of liquid reaching the mouth. Learning requires temporal contiguity between the conditioned stimulus, or movement, and the reinforcer, as well as the more frequent occurrence of the reinforcer in

the presence rather than the absence of the conditioned stimulus (contingency). Analysis of the conditions of learning reveals that rewards that are fully predicted do not contribute to learning [5]. Rather, the acquisition of associative strength of a conditioned stimulus depends on the discrepancy between the maximal associative strength sustained by the reinforcer and the current strength of the predictive stimulus (prediction error) [6] and requires in some situations the uncertainty of the reinforcer [7]. Uncertainty is different from probability; it is highest at probability (p)=0.5 and decreases toward lower and higher probabilities, where reward absence or presence becomes increasingly certain. Uncertainty can be assessed as entropy, variance, and the associability term of attentional learning theories [7].

According to game theory and microeconomics, the value of rewards for behavioural reactions and decisions can be assessed from the multiplicative product of magnitude and probability of the future reward (expected reward value). In addition, the delay to the future reward reduces the reward value hyperbolically [8]. However, the simple products of magnitude, probability and delay do not always explain how individuals value rewards at all probabilities, and the closely related terms 'utility' [9] and 'prospect' [10] allow a better assessment of the influence of rewards on decision-making. Behavioural preference tests serve to establish the subjective values of the different reward objects. In essence, the reward that is most preferred by an individual, and selected when a choice is available, has the highest utility, irrespective of its particular magnitude, probability or delay. Several additional factors contribute to the subjective perception of reward value, such as the individual's history, their personal assets, their emotional state and the way in which the decision problem is posed (framing). Individuals will try to maximize the utility of the outcome and maintain its stability, given the environmental situation and the behaviour of the other players (Nash equilibrium). Thus, schematically, learning theory describes how organisms react to rewards and acquire new reactions, and game theory and microeconomics assess the value of rewards for decision-making. There are similar basic notions and considerable overlaps among these theories, such as associative strength (learning theory) corresponding to value (game theory), and contrast corresponding to utility and framing. Efficient conditioned behaviour according to learning theory can be viewed as a form of game that reaches Nash equilibrium, such as matching behaviour [11], in which players reach optimal returns by adapting their response rates to the relative frequencies of rewards.

Behavioural ecology uses the basic terms of game theory for an evolutionary understanding of reward functions [12,13]. The utility of a reward and its maximisation contribute to fitness for survival. The game for reward

is played against nature and against other players, and the reward value is assessed as payoff by subtracting cost from gain. According to this view, evolutionarily stable behavioural strategies (ESS) evolve that help to maximise the payoffs, and brains are selected that are able to calculate, compare and maximise payoff, maintain ESSs and thus maximise the fitness for survival. Without the evolutionary perspective, animal learning and game theory might be short sighted and prone to local optima but long-term errors.

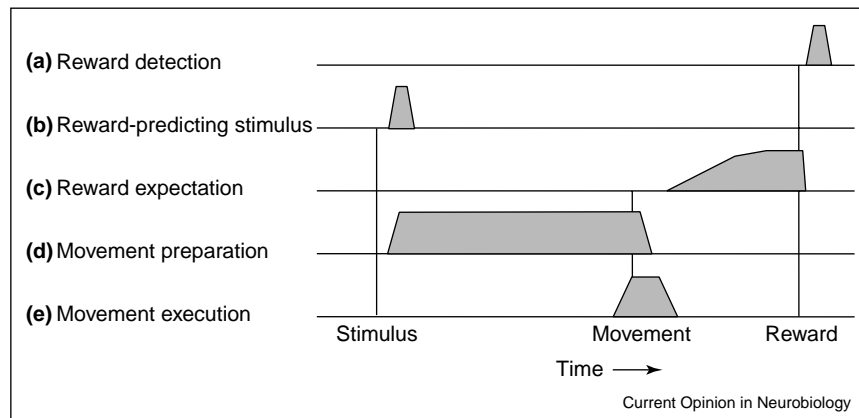
Forms of reward coding

On the basis of learning and game theories, we can conceptualise how individual neurons can process rewards for maximal use by using behavioural tasks that are commonly employed by experimenters for investigating specific brain structures and behavioural processes, such as delayed response tasks for studying frontal cortex and basal ganglia, and Pavlovian learning for assessing reward associations. The reviewed behavioural tasks often give the animal only the choice between correct and incorrect responses, whereas choice tasks need to have a higher degree of freedom by permitting the subject to choose, for example, among different reward objects or among similar rewards occurring at different schedules. Tasks occasionally employ unrewarded single or sequential trials that nevertheless require correct performance, after which the subject can advance to a rewarded trial type. In the simplest form of reward processing neurons detect the receipt of a reward by an increase or decrease in activity after the reward has occurred (Figure 1a). As novel external stimuli are associated with rewards through Pavlovian conditioning, they become reward predictors. Through learning, these reward-predicting stimuli permit subjects to prepare to collect the reward long before it becomes available, thereby providing a substantial advantage over competitors. Some neurons respond to these reward-predicting stimuli (Figure 1b). After the behavioural reaction towards the reward has occurred but before the reward has been consumed, it is useful to maintain a neural representation of the reward, such that attention is maintained and possible interference is avoided. Some neurons maintain elevated activity during this reward expectation period (Figure 1c). When subjects perform behavioural reactions to obtain rewards, neural representations about the action, the reward and their contingency should be maintained during a period in which the movement toward the reward is being prepared and executed (Figures 1d and e). This information would be integrated into the movement-related activity and have an influence on the neural processing that underlies the seeking and acquisition of goal objects.

Coding of basic reward terms

Neurons respond to the receipt of different food and liquid rewards and discriminate among rewards in the

Figure 1



Schematic forms of reward coding. (a–e) show different forms explained in the text. Shaded areas represent activations in different neurons during specific periods of behavioural tasks, reflecting specific forms of reward processing, including influences on movement preparation and execution.

orbitofrontal cortex [14–16], the amygdala [17,18] and the striatum including the nucleus accumbens [19,20]. Orbitofrontal and striatal neurons can discriminate among conditioned stimuli on the basis of the predicted rewards but not on the basis of spatial or object components [15,20]. Other neurons in these structures show differential activations during the expectation of food and liquid rewards following the behavioural reaction [15,20]. Neurons in the dorsolateral prefrontal cortex and the striatum show activations during the mnemonic and movement preparatory periods of delayed response tasks, which reflect differentially the predicted food and liquid rewards [20,21]. When an animal is presented with a reward they prefer more than another, about two-thirds of reward-discriminating orbitofrontal and striatal neurons are strongly activated, whereas the remaining neurons show the opposite relationship and are activated with greater strength by the less preferred reward [15,20].

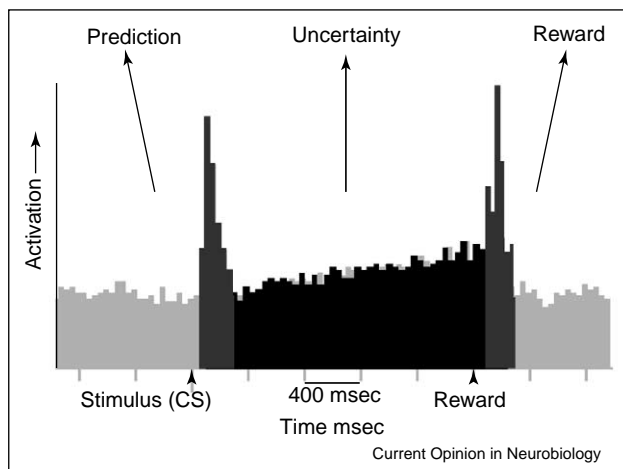
Neurons in several brain structures show stronger task-related activity in all forms of reward-related activations (shown in Figure 1) in rewarded compared to unrewarded trials. The structures include the striatum [22–28], the dorsolateral prefrontal cortex [29,30], the medial prefrontal cortex [29], the orbitofrontal cortex [31], the anterior cingulate cortex [32], the perirhinal cortex [33], the superior colliculus [34], the pars reticulata of substantia nigra [35] and the dopaminergic pars compacta of substantia nigra [36,37]. Inverse reward relationships, with stronger task-related changes in unrewarded rather than rewarded trials, occur in some neurons of the dorsolateral prefrontal cortex [38], the orbitofrontal cortex [31], the striatum [34] and the pars reticulata of substantia nigra [35].

Neurons show stronger task-related activations with higher magnitude of liquid rewards in all reward-related

forms in the striatum [39], the dorsolateral prefrontal cortex [40,41], the orbitofrontal cortex [42,43], the parietal cortex [44,45], the posterior cingulate cortex [46] and the dopaminergic pars compacta of substantia nigra [47,48]. Some neurons in the striatum and in motor regions of the frontal lobe, such as premotor cortex, frontal eye fields and supplementary eye fields, show enhanced activity with increasing reward magnitude, or in the presence compared to absence of reward, which could be related to the movement changes induced by the rewards [41,49]. About one-third of reward-modulated striatal neurons show increasing task-related activations with decreasing reward size [39].

The probability of receiving a reward is a defining factor of reward value according to game theory. Dopamine neurons show increasing phasic responses to conditioned stimuli predicting reward with increasing probability [50,51], and similar increases occur in parietal cortex and striatum during mnemonic and movement preparatory task periods and movements [44,45,52]. However, reward responsive tonically active neurons (TANs) in the striatum do not appear to be sensitive to reward probability [51], which indicates that not all neurons sensitive to reward code its value as defined by game theory. Increasing the expected reward value (magnitude \times probability) increases the task-related activations of parietal neurons [45], and this increase is more important than increasing magnitude or probability separately [53]. In an experiment that involves making choices with varying magnitude and probability, parietal neurons track the recently experienced reward value [54]. Other studies that have investigated probabilistic coding report that neurons in the superior colliculus and the pars reticulata of substantia nigra show stronger target-expecting activity with increasing target probability [55–57].

Figure 2



Two forms of behavior-related activation in dopamine neurons. Phasic prediction error-related responses to rewards ('Reward', dark grey) and reward-predicting stimuli ('Prediction', dark grey) occur in the same neurons as uncertainty-related ramping activity ('Uncertainty', black) during the interval between reward-predicting stimuli and reward. Arrows indicate the onset of the conditioned stimulus (CS, a stimulus that predicts reward) and the time of reward delivery (Reward).

Probabilistic reward tasks also allow the study of uncertainty, which (as explained above) is maximal at $p=0.5$. A proportion of dopamine neurons show activity that increases slowly toward the reward in the interval between a conditioned stimulus and a reward. The ramp is maximal at $p=0.5$, is less prominent at lower and higher probabilities and does not occur when the reward is substituted by a visual stimulus. It could, therefore, reflect reinforcer uncertainty [50]. Most of these dopamine neurons also show phasic responses to reward-predicting stimuli, although the two task modulations are unrelated in occurrence and magnitude in individual neurons (Figure 2). Neurons in the posterior cingulate cortex show increased task-related activations as animals choose rewards from probability distributions with larger variance, which suggests a relationship to risk and uncertainty [46].

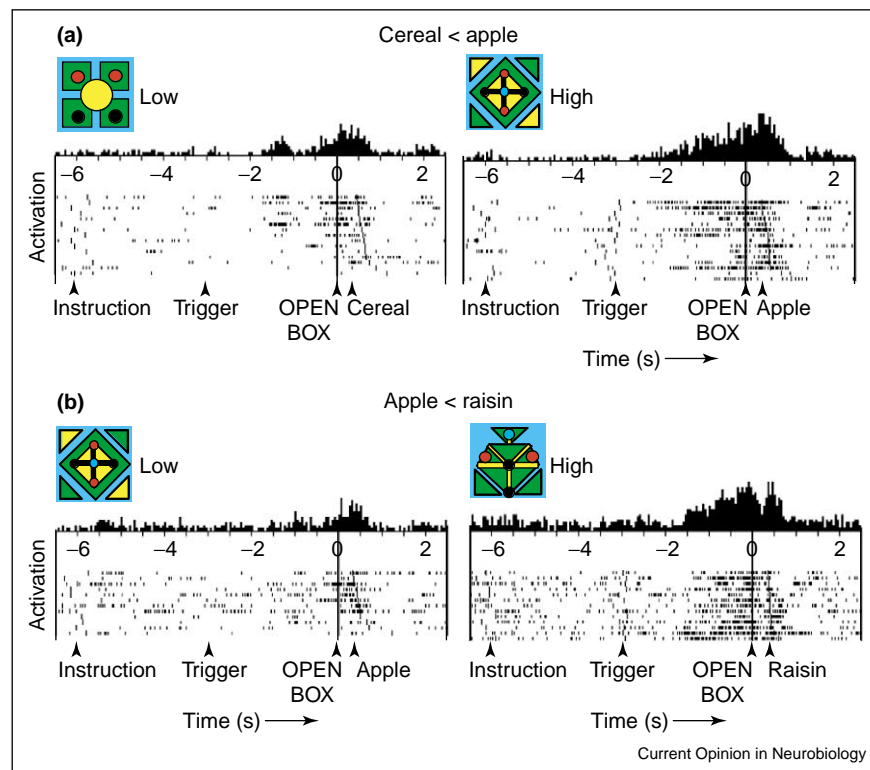
Another factor influencing reward value is delay of reward. The issue of reward delay has not yet been addressed in neurophysiological studies, although the neural coding of time is currently receiving much interest. Activity in parietal neurons reflects the elapsed time of a stimulus, and the time courses of activations follow the temporal distributions of the delay [58,59]. It would be interesting to see how neurons process reward value and reward expectation in such paradigms.

Although neurons in the secondary gustatory area of the posterior orbitofrontal cortex appear to code certain phys-

ical characteristics of rewards, such as the glucose concentration of liquids and the texture of fat [60,61], some neurons in other structures might code the relative, subjective utility of rewards. In a recent experiment, it was found that neurons in the orbitofrontal cortex show higher activation when expecting a small piece of apple than when expecting cereal. However, when in another trial block the cereal is replaced by a piece of raisin that is even more preferred by the animal, the same neurons show higher activity for the raisin than for the apple (Figure 3; [15]). Such neurons appear to be sensitive to the reward that has the highest utility at a given moment for the animal. The difference in neural activity between the lowest and the highest utility increases the gain of coding and thus enhances the ability of organisms to make fine discriminations of utility within a given context. This would be an advantage over a system that operated with linear coding over all reward ranges and without context dependency. In a related experiment, some neurons in the dorsolateral prefrontal cortex show delay-related activations in unrewarded trials that depend on the preference the animal has for the particular food or liquid delivered in the rewarded trials of the same block [38]. The difference in activation in unrewarded trials increases with the difference in value between the rewarded and the unrewarded trials. Apparently, these neurons have a representation of the current reward even in unrewarded trials and indicate the value of reward that the animal is missing. The results indicate that the processing of expected reward utility occurs on the basis of information from rewarded and unrewarded trials. When neurons are investigated during choice behaviour, activity of parietal neurons does not necessarily reflect the expected value (magnitude \times probability), but rather the subjective utility that each option has for the animal in a given trial [53]. Taken together, these experiments reveal a neural correlate for the behavioural reward contrast effect and appear to code the utility of rewarding outcomes in microeconomic terms.

Reward prediction errors usually represent the scalar difference in value (magnitude \times probability) between a delivered and a predicted reward. As some forms of learning progress as a function of prediction error [6,7], investigating the neural coding of prediction error should provide insights into learning mechanisms for short or long-term synaptic modification [62]. Dopamine neurons appear to emit a reward prediction error signal, as they are activated by rewards that are better than predicted, uninfluenced by rewards that occur exactly as predicted and depressed by rewards that are worse than predicted [37,47,48,50,51,62,63]. The activation with positive prediction errors follows a different slope than the depression with negative prediction errors, which could be due to the low background activity of dopamine neurons onto which the depression needs to act [50]. The defining test for the role of prediction errors in learning is the blocking

Figure 3



Relative preference coding for food rewards in a single orbitofrontal neuron. **(a)** The neuron was tested in imperative trials in which a small piece of either cereal or apple was made available after correct task performance or, **(b)** in a separate block of trials, a piece of apple or raisin. The particular reward was indicated at trial onset by a specific instruction picture, shown above the neural histograms and rasters. The neurons showed higher reward expectation-related activity for the preferred reward and thus had the higher utility for the animal (raisin preferred over apple, and apple preferred over cereal). Instruction denotes the time of visual stimulus presentation indicating the target of movement and predicting the kind of reward, trigger denotes stimulus that releases the movement. Dots denote neuronal impulses. Each line of dots shows one trial. In (a) and (b), trials alternated randomly between the two rewards and are separated for analysis. Modified from [15].

paradigm, in which a stimulus that is paired with a fully predicted reward is blocked from learning [6], and the reward responses of dopamine neurons follow the notions of blocking in all respects [36]. Dopamine neurons even code prediction errors when the absence of reward is explicitly predicted in a conditioned inhibition paradigm [64]. Orbitofrontal and striatal neurons respond to unexpected rewards [31,65], and neurons in dorsolateral prefrontal cortex, anterior cingulate, posterior cingulate and frontal eye fields are activated when a reward fails to occur because of an error by the animal [66,67,68], although neural activity coding a bidirectional prediction error as with dopamine neurons is not reported for these structures.

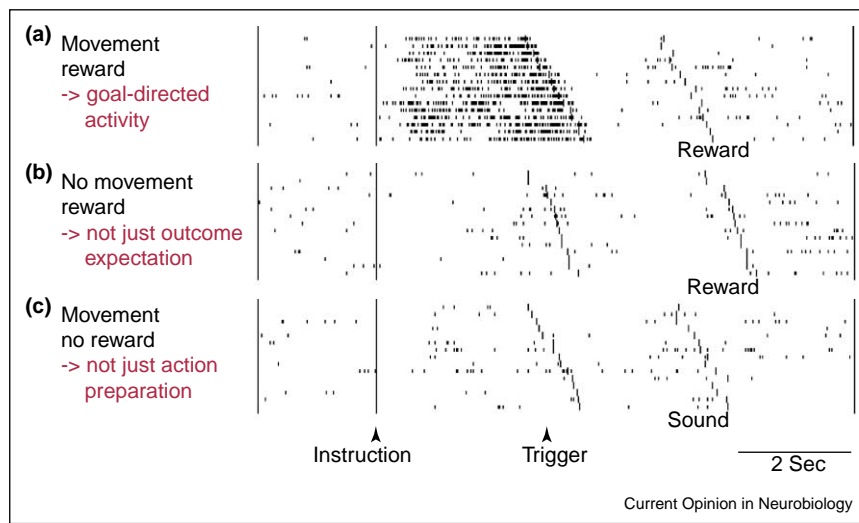
Neural use of reward information

Neural responses to rewards, neural responses to reward-predicting stimuli and activations occurring during the expectation of rewards appear to be involved in the direct detection of reward-related events. However, the neural coding of reward information is not restricted to these

responses. Reward information is present during various task components and influences neural responses to stimuli and activations related to the preparation and execution of arm and eye movements. These influences could play a role in transferring reward information to neural mechanisms involved in stimulus detection and discrimination, decision-making and movement initiation and control. Reward influences on activations related to movements might reflect representations of rewards and action-reward contingencies that are a hallmark of goal directed mechanisms (Figure 4; [1]).

The influence of rewards on neural responses to stimuli and movement-related activations might lead to better signal-to-noise ratios, higher neural discriminations of stimulus and movement parameters, and, thus, higher information content of neural signals. Indeed, neurons in the striatum and the dorsolateral prefrontal cortex show better discriminations (higher differences): during movement preparation between movement and non-movement reactions, during choice of spatial target

Figure 4



Possible neural coding of goal-direction in a single neuron in the caudate nucleus. In each trial type, the neuronal responses of interest are those recorded during the movement preparatory delay period, that is, during a period of time after the instruction to move (or not) had been given, but before the 'go signal' or trigger to make the movement. **(a)** Only in rewarded movement trials (top) was the neural activity during the delay period elevated over baseline. **(b)** Controls indicated no activation in non-movement trials and **(c)** in unrewarded movement trials. Thus, the activation in (a) might represent a neural representation of the reward at the time at which one specific behavioral reaction for obtaining the reward was being prepared. Sound denotes a conditioned reinforcer that signalled correct performance without having intrinsic reward value. Data from [23].

positions between rewarded trials and unrewarded trials [23,25,29] and between preferred rewards and non-preferred rewards [20]. Neurons in the superior colliculus show higher gains of visual responses in rewarded compared to unrewarded trials [34]. Neuronal activations in the dorsolateral prefrontal cortex show increased information content about spatial position in rewarded compared to unrewarded trials (from 0.2 to 0.4 bit in a 1 bit alternative with left-right positions) [30^{**}]. Neurons in parietal cortex are more selective for attended compared to unattended stimuli, as expressed by increased neural signal detection (d'), [68] and show better discrimination of intended reaching targets and higher mutual information with higher reward value (magnitude \times probability) [45]. These data illustrate how the neural processing of reward information could influence neural mechanisms related to behavioural discrimination and goal-directed movement.

One of the key issues in neuroscience concerns the processes that underlie decision-making. When decisions concern goals and goal-directed movements, game theory can become a powerful tool for neuroscience. On the basis of simple variables such as, magnitude, probability and utility of reward, game theory has helped us to understand behavioural decision processes made by individuals. The neural coding of basic game theory variables would provide the inputs to decision-making mechanisms. This might involve, to some extent, Pavlovian attribution of reward value to environmental stimuli.

However, straightforward reinforcement learning does not produce optimal learning and performance in many games, and more refined and heterogeneous information and learning mechanisms are necessary, such as experience-weighted attraction learning [69]. Neurons in the parietal cortex monitor the recently experienced reward value in tasks in which animals match the frequency of behavioural reactions to available reward frequencies [54]. Such activity could constitute an input mechanism for decision-making. The neural coding of relative reward preference and utility in imperative trials [15] and in a decision-making context [53] could in a similar way provide input information for basic decision processes. The next question to be asked is whether or not there are forms of neural activity beyond tracking reward value or utility that are closer to the decision process itself. That type of activity should occur specifically in decision situations and not in imperative trials without much choice [70]. Studies on perceptual decision-making and adaptation to changing instructions reveal a diffusion-like race process in which neural activity rises to a threshold beyond which the behavioural choice occurs [71,72,73^{**},74,75]. The race model relies on continuously incoming information that brings the choice process closer to a decision point. A similar mechanism might also apply to certain reward-related decisions in which reward values evolve slowly and approach a decision threshold, although that mechanism might not necessarily be appropriate for all games.

Conclusions

As rewards are defined by their influence on behaviour, the neural processing of reward information should include a behavioural perspective. The basic parameters of reward coding can be understood on the basis of concepts from animal learning theory, game theory and microeconomics. We are beginning to unravel the way in which the brain extracts reward information from multimodal stimuli and the way in which these parameters are coded in different brain structures. In addition, we might be on the way to understanding the brain processes that are engaged when we make priority decisions based on an evaluation of reward values. It would help to have more general knowledge about the neural mechanisms that underlie basic decision processes in order to understand reward-directed decisions. At the same time, because reward-directed decisions are such a basic part of decision-making, their study might be useful for a general understanding of general neural decision-making processes.

Acknowledgements

This work would not have been possible without the contributions of HC Cromwell, A Dickinson, CD Fiorillo, OK Hassani, JR Hollerman, PN Tobler, L Tremblay and P Waelti. The author's writing and experiments were supported by the Wellcome Trust and the Swiss National Science Foundation.

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. Dickinson A, Balleine B: **Motivational control of goal-directed action.** *Anim Learn Behav* 1994, **22**:1-18.
 2. Flaherty CF: *Incentive relativity.* Cambridge, UK: Cambridge University Press; 1996.
 3. Brosnan SF, de Waal FBM: **Monkeys reject unequal pay.** *Nature* 2003, **425**:297-299.
 4. Wise RA: **Brain reward circuitry: insights from unsensed incentives.** *Neuron* 2002, **36**:229-240.
 5. Kamin LJ: **Selective association and conditioning.** In *Fundamental Issues in Instrumental Learning.* Edited by Mackintosh NJ, Honig WK. Dalhousie, Canada: Dalhousie University Press; 1969: 42-64.
 6. 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, USA: Appleton Century Crofts; 1972: 64-99.
 7. Pearce JM, Hall G: **A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli.** *Psychol Rev* 1980, **87**:532-552.
 8. Ho MY, Mobini S, Chinang T-J, Bradshaw CM, Szabadi E: **Theory and method in the quantitative analysis of 'impulsive choice' behaviour: implications for psychopharmacology.** *Psychopharmacology* 1999, **146**:362-372.
 9. Bernoulli J: **Exposition of a new theory on the measurement of risk.** *Econometrica* 1954, **22**:23-36.
 10. Kahneman D, Tversky A: **Choices, values, and frames.** *Am Psychol* 1984, **4**:341-350.
 11. Herrnstein RJ: **Relative and absolute strength of response as a function of frequency of reinforcement.** *J Exp Anal Behav* 1961, **4**:267-272.
 12. Lewontin RC: **Evolution and the theory of games.** *J Theor Biol* 1961, **1**:382-403.
 13. Maynard Smith J: *Evolution and the Theory of Games.* Cambridge, UK: Cambridge University Press; 1982.
 14. Thorpe SJ, Rolls ET, Maddison S: **The orbitofrontal cortex: neuronal activity in the behaving monkey.** *Exp Brain Res* 1983, **49**:93-115.
 15. Tremblay L, Schultz W: **Relative reward preference in primate orbitofrontal cortex.** *Nature* 1999, **398**:704-708.
 16. Hikosaka K, Watanabe M: **Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards.** *Cereb Cortex* 2000, **10**:263-271.
 17. Nishijo H, Ono T, Nishino H: **Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance.** *J Neurosci* 1988, **8**:3570-3583.
 18. Nakamura K, Mikami A, Kubota K: **Activity of single neurons in the monkey amygdala during performance of a visual discrimination task.** *J Neurophysiol* 1992, **67**:1447-1463.
 19. Bowman EM, Aigner TG, Richmond BJ: **Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards.** *J Neurophysiol* 1996, **75**:1061-1073.
 20. Hassani OK, Cromwell HC, Schultz W: **Influence of expectation of different rewards on behavior-related neuronal activity in the striatum.** *J Neurophysiol* 2001, **85**:2477-2489.
 21. Watanabe M: **Reward expectancy in primate prefrontal neurons.** *Nature* 1996, **382**:629-632.
 22. Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M: **Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning.** *J Neurosci* 1994, **14**:3969-3984.
 23. Hollerman JR, Tremblay L, Schultz W: **Influence of reward expectation on behavior-related neuronal activity in primate striatum.** *J Neurophysiol* 1998, **80**:947-963.
 24. Shidara M, Aigner TG, Richmond BJ: **Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials.** *J Neurosci* 1998, **18**:2613-2625.
 25. Kawagoe R, Takikawa Y, Hikosaka O: **Expectation of reward modulates cognitive signals in the basal ganglia.** *Nat Neurosci* 1998, **1**:411-416.
 26. Watanabe K, Lauwereyns J, Hikosaka O: **Neural correlates of rewarded eye movements in the primate caudate nucleus.** *J Neurosci* 2003, **23**:10052-10057.
 27. Ravel S, Legallet E, Apicella P: **Tonically active neurons in the monkey striatum do not preferentially respond to appetitive stimuli.** *Exp Brain Res* 1999, **128**:531-534.
 28. Shimo Y, Hikosaka O: **Role of tonically active neurons in primate caudate in reward-oriented saccadic eye movement.** *J Neurosci* 2001, **21**:7804-7814.
 29. Matsumoto K, Suzuki W, Tanaka K: **Neuronal correlates of goal-based motor selection in the prefrontal cortex.** *Science* 2003, **301**:229-232.
 30. Kobayashi S, Lauwereyns J, Koizumi M, Sakagami M, Hikosaka O: **Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex.** *J Neurophysiol* 2002, **87**:1488-1498.
- After the pioneering study by Watanabe [21] on reward-discriminating prefrontal activity, this work makes the next step by investigating the actual gain in neural information derived from the expectation of reward. Application of basic information theory reveals this activity as a potential input for decision-making mechanisms, although it is presently unclear how close the activity actually is to a neural decision-making mechanism.
31. Tremblay L, Schultz W: **Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex.** *J Neurophysiol* 2000, **83**:1864-1876.

32. Shidara M, Richmond BJ: **Anterior cingulate: single neuron •• signals related to degree of reward expectancy.** *Science* 2002, **296**:1709-1711.

After neural activity related to reward expectation has been observed in several brain structures, this study demonstrates a long-term activity that anticipates the reward and progressively increases over several trials through intervening stimuli and movements, indicating a very stable reward representation.

33. Liu Z, Richmond BJ: **Response differences in monkey TE and perirhinal cortex: stimulus association related to reward schedules.** *J Neurophysiol* 2000, **83**:1677-1692.
34. Ikeda T, Hikosaka O: **Reward-dependent gain and bias of visual responses in primate superior colliculus.** *Neuron* 2003, **39**:693-700.
35. Sato M, Hikosaka O: **Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement.** *J Neurosci* 2002, **22**:2363-2373.
36. Waelti P, Dickinson A, Schultz W: **Dopamine responses comply with basic assumptions of formal learning theory.** *Nature* 2001, **412**:43-48.
37. Kawagoe R, Takikawa Y, Hikosaka O: **Reward-predicting activity of dopamine and caudate neurons – a possible mechanism of motivational control of saccadic eye movement.** *J Neurophysiol* 2004, **91**:1013-1024.
38. Watanabe M, Hikosaka K, Sakagami M, Shirakawa SI: **Coding and •• monitoring of behavioral context in the primate prefrontal cortex.** *J Neurosci* 2002, **22**:2391-2400.
- This ingenious study reports that neurons in dorsolateral prefrontal cortex show differential delay activations in unrewarded trials that depend on which reward is given in the other rewarded trials within the same trial block, indicating what the animal is actually missing in the current trial.
39. Cromwell HC, Schultz W: **Effects of expectations for different reward magnitudes on neuronal activity in primate striatum.** *J Neurophysiol* 2003, **89**:2823-2838.
40. Leon MI, Shadlen MN: **Effect of expected reward magnitude on the responses of neurons in the dorsolateral prefrontal cortex of the macaque.** *Neuron* 1999, **24**:415-425.
41. Roesch MR, Olson CR: **Impact of expected reward on neuronal • activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex.** *J Neurophysiol* 2003, **90**:1766-1789.
- The authors present a huge effort to separate the potential coding of reward value from the changes of movement-related activity induced by rewards.
42. Wallis JD, Miller EK: **Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task.** *Eur J Neurosci* 2003, **18**:2069-2081.
43. Roesch MR, Olson CR: **Distinguishing reward representation from motivational modulation in primate frontal cortex** [abstract]. *Society for Neuroscience*; 2003: 724.2.
44. Platt ML, Glimcher PW: **Neural correlates of decision variables in parietal cortex.** *Nature* 1999, **400**:233-238.
45. Corneil BD, Musallam S, Andersen RA: **Representation of reward expectancy in the medial bank of the intraparietal sulcus: implications for neural prosthetics** [abstract]. *Society for Neuroscience*; 2003: 607.8.
46. McCoy AN, Platt ML: **Reward signals in posterior cingulate cortex** [abstract]. *Society for Neuroscience*; 2003: 385.6.
47. Satoh T, Nakai S, Sato T, Kimura M: **Correlated coding of motivation and outcome of decision by dopamine neurons.** *J Neurosci* 2003, **23**:9913-9923.
48. Bayer HM, Glimcher PW: **Midbrain dopamine neurons encode a quantitative reward prediction error signal** [abstract]. *Society for Neuroscience*; 2003: 557.9.
49. Lauwereyns J, Watanabe K, Coe B, Hikosaka O: **A neural correlate of response bias in monkey caudate nucleus.** *Nature* 2002, **418**:413-417.
50. Fiorillo CD, Tobler PN, Schultz W: **Discrete coding of reward probability and uncertainty by dopamine neurons.** *Science* 2003, **299**:1898-1902.
51. Morris G, Arkadir D, Nevet A, Bergman H: **Basal ganglia neuronal teachers: cooperating different messages** [abstract]. *Society for Neuroscience*; 2003: 704.12.
52. Samejima K, Ueda Y, Kenji D, Minoru K: **Activity of striate projection neurons encodes action - selective reward expectations** [abstract]. *Society for Neuroscience*; 2003: 704.9.
53. Dorris MC, Glimcher PW: **Lateral intraparietal area activity during manipulations of the expected value and utility of choices** [abstract]. *Society for Neuroscience*; 2003: 767.1.
54. Sugrue LP, Rorie AE, Corrado GS, Newsome WT: **Matching behavior is effectively modeled by a simple linear integrator of rewards** [abstract]. *Society for Neuroscience*; 2003: 385.1.
55. Basso MA, Wurtz RH: **Modulation of neuronal activity by target uncertainty.** *Nature* 1997, **389**:66-69.
56. Dorris MC, Munoz DP: **Saccadic probability influences motor preparation signals and time to saccadic initiation.** *J Neurosci* 1998, **18**:7015-7026.
57. Basso MA, Wurtz RH: **Neuronal activity in substantia nigra pars reticulata during target selection.** *J Neurosci* 2002, **22**:1883-1894.
58. Leon MI, Shadlen MN: **Representation of time by neurons in the • posterior parietal cortex of the macaque.** *Neuron* 2003, **38**:317-327.
- The authors present the first modern study on the coding of time by the brain, which is one of the great enigmas of neuroscience. Some preliminary work on this topic was done in the 1970s by other investigators. The work is also interesting from the temporal aspect of decision-making (when to make a behavioural response).
59. Janssen P, Shadlen MN: **Representation of the hazard function of elapsed time by neurons in macaque area LIP** [abstract]. *Society for Neuroscience*; 2003: 767.2.
60. Rolls ET, Yaxley S, Sienkiewicz ZJ: **Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey.** *J Neurophysiol* 1990, **64**:1055-1066.
61. Rolls ET, Critchley HD, Browning AS, Hernadi I, Lenard L: **Responses to the sensory properties of fat of neurons in the primate orbitofrontal cortex.** *J Neurosci* 1999, **19**:1532-1540.
62. Schultz W: **Predictive reward signal of dopamine neurons.** *J Neurophysiol* 1998, **80**:1-27.
63. Schultz W, Dayan P, Montague RR: **A neural substrate of prediction and reward.** *Science* 1997, **275**:1593-1599.
64. Tobler PN, Dickinson A, Schultz W: **Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm.** *J Neurosci* 2003, **23**:10402-10410.
65. Apicella P, Legallet E, Trouche E: **Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states.** *Exp Brain Res* 1997, **116**:456-466.
66. Watanabe M: **The appropriateness of behavioral responses coded in post-trial activity of primate prefrontal units.** *Neurosci Lett* 1989, **101**:113-117.
67. Ito S, Stuphorn V, Brown JW, Schall JD: **Performance monitoring • by the anterior cingulate cortex during saccade countermanding.** *Science* 2003, **302**:120-122.
- This study follows up on an earlier issue of error coding in the frontal cortex [66]. The authors examine reward prediction errors in an interesting task in which monkeys are sometimes asked to withhold an already prepared saccadic eye movement with different levels of difficulty.
68. Bendiksy MS, Platt ML: **Motivation focuses attention and enhances neuronal selectivity in parietal cortex** [abstract]. *Society for Neuroscience*; 2003: 385.5.
69. Camerer CF, Ho TH: **Experience-weighted attraction learning in normal form games.** *Econometrica* 1999, **67**:827-874.
70. Barraclough DJ, Conroy ML, Lee D: **Conjunctive coding of decision variables in prefrontal cortex during a competitive game** [abstract]. *Society for Neuroscience*; 2003: 722.21.
71. Newsome WT, Britten KH, Movshon JA: **Neuronal correlates of a perceptual decision.** *Nature* 1989, **341**:52-54.

72. Kim JN, Shadlen MN: **Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque.** *Nat Neurosci* 1999, **2**:176-185.

73. Gold JI, Shadlen MN: **Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward.** *Neuron* 2002, **36**:299-308.

An interesting charming and mathematically illustrated description of how the development of code-deciphering mechanisms during WW2 by Alan Turing and colleagues could lead to a general description of decision

mechanisms that resemble neural mechanisms involved in decision-making with continuously incoming information.

74. Schall JD, Stuphorn V, Brown JW: **Monitoring and control of action by the frontal lobes.** *Neuron* 2002, **36**:309-322.

75. Romo R, Hernandez A, Zainos A, Lemus L, Brody CD: **Neuronal correlates of decision-making in secondary somatosensory cortex.** *Nat Neurosci* 2002, **5**:1217-1225.