CHOOSING THE GREATER OF TWO GOODS: NEURAL CURRENCIES FOR VALUATION AND DECISION MAKING

Leo P. Sugrue, Greg S. Corrado and William T. Newsome

Abstract | To make adaptive decisions, animals must evaluate the costs and benefits of available options. The nascent field of neuroeconomics has set itself the ambitious goal of understanding the brain mechanisms that are responsible for these evaluative processes. A series of recent neurophysiological studies in monkeys has begun to address this challenge using novel methods to manipulate and measure an animal's internal valuation of competing alternatives. By emphasizing the behavioural mechanisms and neural signals that mediate decision making under conditions of uncertainty, these studies might lay the foundation for an emerging neurobiology of choice behaviour.

> "My diving bell becomes less oppressive, and my mind takes flight like a butterfly. There is so much to do. You can wander off in space or in time, set out for Tierra del Fuego or for King Midas's court. You can visit the woman you love, slide down beside her and stroke her still-sleeping face. You can build castles in Spain, steal the Golden Fleece, discover Atlantis, realize your childhood dreams and adult ambitions."[1]

In 1995, a brainstem stroke left French magazine editor Jean-Dominique Bauby 'locked in', robbed of all voluntary movement save the ability to blink his left eyelid. The sentences above are excerpted from his memoir, a work that he dictated word by word, blinking to select each letter as the alphabet was recited to him over and over. His words are a powerful reminder that decisions, the ultimate expressions of will, can be dissociated from the actions through which they are commonly manifest, and owe their true existence to processes hidden within the recesses of the brain. For centuries, students of philosophy, psychology and behavioural economics have sought access to this inner world through introspection and the study of overt behaviour. More recently, neuroscientists have joined their ranks, eager to understand the mechanisms of decision making at the cellular and circuit level.

This shared interest in decision making has resulted in the emergence of a new interdisciplinary field of research, often referred to as 'neuroeconomics', the expressed goal of which is to understand the neural basis of individual choice behaviour[2]. This field encompasses behavioural, imaging and physiological approaches in both humans and animals. In this review, we discuss data only from electrophysiological experiments, primarily recordings from single neurons in behaving monkeys. By focusing on one of the more developed lines of research in this field, we hope to move beyond the particular to illuminate some challenges and approaches that are relevant to the field in general.

The form of our review recapitulates the development of this topic. We begin by summarizing the progress that has been made during the past decade in uncovering the neural correlates of simple perceptual decisions. We then introduce our primary focus — the interaction between decision making and reward — by exploring some key concepts that relate to the neurobiology of reward and motivation. Finally, we turn to recent attempts that have been made to expand the neurobiology of decision making to account for the powerful influences of reward history and expectation on choice behaviour.

Howard Hughes Medical Institute and Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA. Correspondence to L.P.S. and W.T.N.
e-mail: leo@monkeybiz.stanford.edu; bill@monkeybiz.stanford.edu
doi:10.1038/nrn1666
Published online 15 April 2005

© 2005 Nature Publishing Group

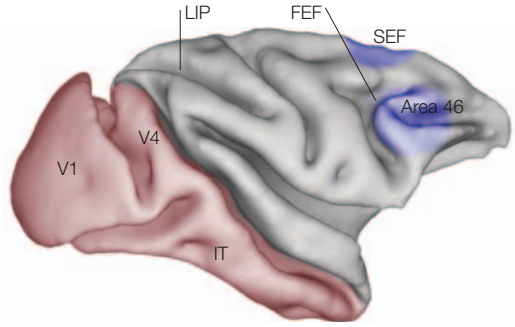

Figure 1 | **Visual and oculomotor systems of the primate brain.** Lateral view of the cerebral hemisphere of the macaque monkey showing visual (red) and visuo-oculomotor (blue) areas. Example visual areas include primary visual cortex (V1), fourth visual area (V4) and inferotemporal cortex (IT). Visuo-oculomotor areas include the lateral intraparietal area (LIP), frontal eye field (FEF), supplementary eye field (SEF) and Walker's cytoarchitectonic area 46 (area 46). Another important oculomotor structure, the superior colliculus, is located in the midbrain, which is not visible on this image. The map of cortical areas was derived from a macaque atlas data set[106], which was downloaded and processed using Caret software[107]. Adapted, with permission, from REF. 106 © (2004) MIT Press, and from REF. 107 © (2001) JAMA.

Neural correlates of perceptual decisions

The neurophysiology of decision making owes its existence to progress in delineating the sensory[3] and motor[4–6] systems of the cerebral cortex. As sensory and motor physiologists traced input and output pathways deeper into the brain, they often found themselves studying the same central brain structures from different perspectives. Because decision making forms a key link between sensation and action, some of these investigators wondered whether the neural transformations responsible for decision making might be implemented in these sensorimotor watershed areas[7–12].

Most, although not all[12–14], efforts to understand these transformations have drawn on the visual and oculomotor systems of the primate brain. Primary visual processing occurs within a network of areas in the occipital and temporal cortices (FIG. 1; red), whereas visuomotor areas that link these sensory areas to the brainstem oculomotor nuclei reside in the frontal and parietal cortices and in the midbrain (FIG. 1; blue). In the context of a typical two-alternative forced-choice discrimination task, these intermediate areas must implement at least three fundamental transformations[15] (FIG. 2a). First, a 'sensory transformation' generates a higher-order visual representation from primary visual input. Second, a 'decision transformation' maps this sensory evidence onto the probability of one or other operant response. Finally, a process of action implementation renders a discrete behavioural response from this probabilistic representation.

These transformations can be studied in monkeys that have been trained to discriminate noisy visual stimuli and report their perceptual judgements using eye movements. By recording neural activity in visuomotor pathways during such tasks, investigators have identified sensory representations as well as decision-related signals in areas of the parietal and frontal cortices. At the neural level, differentiating sensory signals from decision-related signals is relatively straightforward. First, sensory signals require the presence of the sensory stimulus, and extinguish with stimulus offset. Second, and more importantly, in discrimination tasks in which behavioural decisions and neural activity are measured across a range of stimulus strengths, animals make both correct and incorrect judgements in response to the presentation of identical stimuli. On these trials, sensory neurons encode the visual stimulus itself, whereas the activity of decision-related neurons reflects the animal's ultimate choice.

It is more difficult to dissociate decision-related activity from premotor signals, which might reflect the outcome of a decision but not the process of deliberation itself. To distinguish these alternatives, we must look at factors that affect decision making but that are not necessarily expressed in the motor response. For example, decisions vary in their certainty: subjects are very certain about decisions that relate to suprathreshold stimuli but are much less certain when stimuli are at, or below, the psychophysical threshold. By contrast, the motor responses that subjects use to indicate their choices can be highly stereotyped and largely independent of the strength of the evidence on which the underlying decision was based.

Shadlen and Newsome[16,17] showed this dissociation in monkeys that had been trained to discriminate the direction of coherent motion in a stochastic random dot display (FIG. 3a). As the strength of the motion signal was varied from weak (most dots moved in random directions) to strong (most dots moved coherently in a single direction), the animal's behaviour varied between chance and perfect performance (FIG. 3b). Neural signals in the lateral intraparietal area (LIP) covaried with the animal's final decision, but were also modulated by the quality of the sensory evidence (motion coherence). Importantly, this correlation was independent of any covariation between stimulus coherence and the details of the animal's motor response. For the same task, similar decision-related signals have been seen in the dorsolateral prefrontal cortex (DLPFC)[18] and in the superior colliculus[19], a brainstem target of these cortical areas. These results have prompted the proposal that LIP, in cooperation with other areas, implements the decision transformation in this task, converting a sensory representation of visual motion into a decision variable that is capable of guiding behaviour. Additional support for this proposal comes from recent demonstrations that the profile of LIP activity in a reaction time version of this task predicts not only what the animal decides, but also when that decision has been reached[20], and that electrical microstimulation of area LIP influences the outcome of the decision process, as predicted by the properties of the neurons at the stimulation site[21].

Sensorimotor areas like LIP access not only visual input, but also other types of information that are pertinent to decision making. For example, the activity of LIP neurons can be modulated by the attentional priority

© 2005 Nature Publishing Group

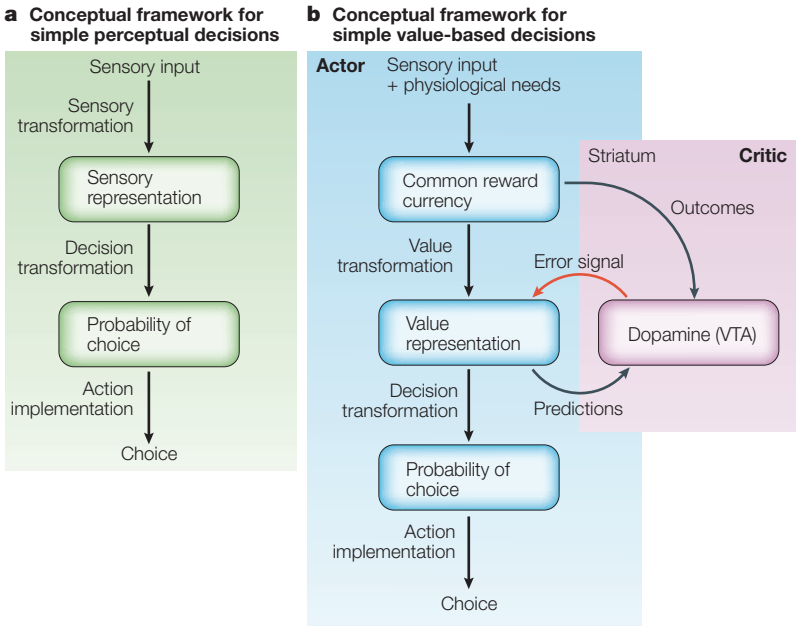

Figure 2 | **Conceptual frameworks for decision making.** A conceptual framework that illustrates proposed processing stages for the formation of simple perceptual and value-based decisions. **a** | Perceptual decisions. A sensory transformation operates on primary sensory input to generate a representation of a higher-order stimulus dimension (for example, visual motion or auditory space). A decision transformation maps this sensory representation onto the probability of alternative operant responses. A final processing stage renders the actual binary decision, reducing the continuous probabilistic representation to a discrete plan for motor action. **b** | Value-based decisions. The absence of a dedicated sensory system for transducing rewards means that sensory input and physiological needs must first interact to identify 'rewards' in the animal's environment. For simplicity, we assume that this initial processing produces a 'common reward currency', which can be considered as the primary input to subsequent stages. This framework is considered to have an 'actor–critic' architecture. Within the actor component, the reward input is transformed into a higher-order representation of the value of different stimuli. Through the action of the critic, this mapping can be optimized to the environment. A decision transformation maps this value representation onto the probability of available behavioural responses. At the final stage, this representation is reduced to a single behavioural choice. VTA, ventral tegmental area, the midbrain origin of the dopaminergic neurons that contribute the 'error signal' to our proposed actor–critic architecture.

of spatial locations[22], the prior probability that a particular eye movement will be instructed[23] and the magnitude of a juice reward associated with a motor response[23]. So, LIP seems to encode a mixture of sensory, motor and cognitive signals that might guide decisions about upcoming behavioural responses (BOX 1).

Decision making and reward

Making adaptive decisions. Consider a fisherman standing on a riverbank deciding where to cast his line. Unable to see the fish, he must instead 'read the water', observing where it slows into deeper, shady pools — places where trout are more likely to feed. As a frequent visitor to this spot, he has had more success at some of these locations than at others. Like many everyday decisions, the correct choice in this situation is not immediately evident from sensory input, but emerges from analysing that input in the context of the value/utility of available stimuli and actions — in this case, our fisherman's accumulated experience and expectation of the likely abundance of fish. Psychologists and economists have long appreciated the influence of reward and valuation on decision making in higher mammals[24], but these factors were notably absent from our preceding discussion. Although reward is an implicit variable in every operant task, most physiological studies of perceptual decision making hold reward contingencies constant to isolate activity that is specifically related to sensorimotor transformations (FIG. 2a). Only recently have investigators begun to manipulate reward independently in order to explore the neural basis of valuation and adaptive behaviour.

A conceptual framework within which to consider value-based choice is proposed in FIG. 2b. Neither conclusive nor complete, it is intended as a starting point from which to discuss the basic steps in building an internal representation of value and using that representation to guide behaviour. Focus first on the left-hand side of this diagram (labelled 'actor'). Like the proposed framework for perceptual decisions (FIG 2a), this framework for value-based decision making comprises three key processing stages. At the first stage, a value transformation takes the input — rewards, or 'common reward currency' — to the system and abstracts from it a representation of the value of available options. At the second stage, a decision transformation maps this value representation onto the probability of alternative courses of action. A final processing stage transforms this continuous probability into a discrete choice among these alternatives. Much of our discussion of the neurobiology of value-based choice focuses on the neural implementation of these transformations; however, the first question raised by this framework is the nature of the input to the system — the rewards themselves.

A common neural currency for rewards. Operationally, we can define a reward as anything that an animal will work to acquire. This definition emphasizes the motivational ('wanting') rather than the affective ('liking') dimension of reward. Both of these dimensions are thought to have distinct psychological and physiological substrates[25]. No single sense organ is responsible for transducing rewards, but the phenomenon of brain-stimulation reward[26] (BSR) indicates that there is a dedicated network of neural structures devoted to reward processing. In the rat, Shizgal and colleagues[27] have shown that BSR can summate with and be traded off against various natural rewards to influence choice behaviour, and that manipulations of physiological state that profoundly modify the value of natural rewards have little influence on the efficacy of BSR. These results indicate that BSR contributes something like a pure reward signal to the neural circuitry that is responsible for valuation. This circuitry interacts with natural rewards downstream from the point where they have been scaled for their identifying features and projected onto a single common dimension of reward. For simplicity, FIG. 2b assumes the existence of such a common reward signal as input to the valuation system.

© 2005 Nature Publishing Group

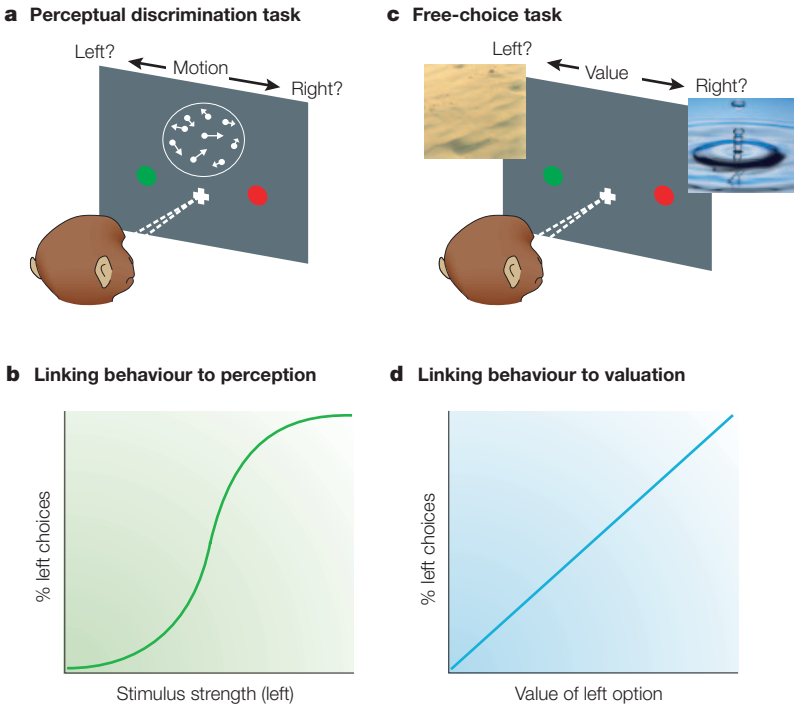

Figure 3 | **Decision-making tasks. a** | General structure of a perceptual discrimination task, in which a monkey reports its judgement of the direction of motion in a random dot stimulus with an eye movement to one of two targets that are aligned with the axis of stimulus motion. **b** | In a perceptual discrimination task, a psychometric function quantifies the reliance of behaviour on the sensory signals being manipulated experimentally. **c** | The general structure of a free-choice task, in which a monkey uses eye movements to indicate its relative valuation of two competing choice targets that are associated with independent underlying reward schedules. The visual icons depict success (water drop) or failure (sand) in acquiring a liquid reward. **d** | In a free-choice task, a valumetric function quantifies the reliance of behaviour on the value signals being manipulated experimentally.

Incentives and errors. BSR is elicited most reliably at sites defined by the projections of midbrain dopamine neurons. Various clinical[28–30] and experimental[31] data also indicate a central role for the dopamine system in processing the motivational aspect of reward. More recent experimental work is beginning to clarify this role[32–34]. In particular, electrophysiological recordings of midbrain dopamine neurons in alert monkeys during simple conditioning tasks[35–37] indicate that dopamine neurons do not signal the occurrence of rewards *per se*. Instead, phasic changes in the firing rates of these cells seem to comprise an error signal, the magnitude and sign of which encode disparities between an animal's expectations and experience of reward.

Because such error signals indicate a flaw in the animal's understanding of the world, they are particularly useful for guiding learning[38]. In what are referred to as 'actor–critic models'[39], an error signal such as that conveyed by dopamine neurons can be used to select actions that maximize rewards in response to specific sensory cues[40,41]. This approach has been successful in accounting for the behaviour of relatively simple organisms in restricted environments[42]. In mammals, the action of dopamine on striatal circuits has been proposed to implement such an actor–critic architecture[43], and recordings from monkey caudate neurons during simple associative conditioning tasks show activity that is consistent with the creation of such stimulus–response bonds[44–46]. However, the direct yoking of stimuli to actions and outcomes implied by the current generation of these models fails to capture the facility with which higher organisms construct complex representations of value and flexibly link them to action selection (BOX 2).

Responding to these limitations, more recent theoretical proposals have expanded the role of the dopamine signal to include the shaping of more abstract models of valuation[47–49]. Consistent with this approach, FIG. 2b portrays the dopamine system as a critic whose influence extends beyond the generation of simple associative predictions to the construction and modification of complex value transformations. In this scheme, the striatum is considered to have the crucial role of liaison between actor and critic. If correct, this proposal indicates that dopamine neurons have access to the value representation depicted in FIG. 2b. Consistent with this idea, Nakahara and colleagues[50] recently showed that dopamine responses were strongly modulated by contextual information that pertained to the evolution of reward probability across successive trials in a task, even when this information was not accompanied by any explicit sensory cue. Further experiments involving more complex value-based tasks will be needed to determine the true scope of the dopamine signal in shaping value-based choice.

Value-based decision making

The cortex as the stage for valuation. The cortex is the likely stage on which value-based decisions unfold. Anatomically, several regions within the prefrontal and parietal association cortices are positioned to link rewards to behavioural responses. These areas receive sensory inputs, project to motor planning centres, and are reciprocally connected to the dopamine system, either directly or through the striatum. Many of these areas show activity that is modulated by the expectation or delivery of reward[23,51–62], or that persists through time in tasks that engage short-term memory[63–69] — properties that indicate a basis for maintaining a representation of value through time. Therefore, whereas lesion data[70,71] have tended to emphasize the importance of the prefrontal cortex in monitoring the motivational value of stimuli and actions[47,49,72,73], physiological studies indicate that this function might, in fact, be distributed across both frontal and parietal cortices[23,74].

In FIG. 2b, the process of interpreting input in the context of behavioural performance is indicated by the 'value transformation', and it produces a value representation — an internal model of the world in which stimuli or actions are scaled for their value or utility to the animal. Developed through the influence of the critic, the mapping from experience to value that creates this representation should be flexible, capable of generalizing across stimuli and able to integrate experience over

© 2005 Nature Publishing Group

Box 1 | **Towards a unified theory of lateral intraparietal function**

A recent opinion piece[90] cautions investigators against rushing to attach specific psychological labels to particular neural signals. Over the past generation, cognitive science has matured beyond strict behaviourist constructions, but our continued inability to cleanly dissociate concepts like 'attention' and 'reward expectancy' indicates that we should continue to heed the behaviourists' admonition to focus on quantifiable variables that yield explanatory power in accounting for animal choice. Descriptively, the scope of these variables might seem limited, as they are necessarily framed in terms of information that is relevant to behaviour in a particular task — the strength of a noisy sensory signal in a discrimination task or the history of choices and rewards in a foraging game — but they have the important advantage that their validity can be quantitatively tested and compared.

A single broad conceptual framework might one day unite the discoveries made through this operational approach. For example, in the context of two-alternative forced-choice eye movement tasks, it has been proposed that lateral intraparietal (LIP) activity reflects a general decision variable that is monotonically related to the logarithm of the likelihood ratio that the animal will select one of the two available eye movements[10,85]. This formalism is appealing both for its mathematical precision and because it suggests how multiple sources of information might be combined into a final decision variable by simply summing the likelihood ratios derived from various inputs — an operation that neurons perform well. Because it views LIP activity as a 'common currency' for integrating information that is pertinent to the selection of future shifts in gaze or attention, this proposal has obvious connections to the even more general concept of a 'saliency map'. In psychophysics and computational vision, saliency maps are conceived as important central clearing houses that prioritize the deployment of limited visual processing resources[91,92], and saliency has already been suggested as a unifying explanation for LIP activity[22].

We believe that it is far too early to predict whether a single broad concept such as saliency will find a mathematically precise and biophysically plausible implementation, and, ultimately, provide a unifying theory of LIP function. Fortunately, data needed to help us make this determination require exactly the same experimental approach advocated above: experiments designed to link specific cognitive phenomena to behaviour that explore these links through quantitative behavioural modelling, and that leverage the insight and variables derived from these models to search for underlying neural signals.

time. Such an internal model frees the organism from slavish adherence to the immediate consequence of every encounter and action, allowing it to organize goal-directed behaviour on a level and time scale that is appropriate for its environment.

Understanding the neural basis of value-based choice entails investigating the neural substrates of these value representations and the mechanisms by which they are constructed, maintained and transformed into behaviour at the decision stage. Most studies that have examined reward-related activity have done so in the context of imperative tasks in which an animal's behavioural response is instructed unambiguously on every trial[75]. This approach is fundamentally limited, because the value transformation in such tasks is rudimentary (the probability of reward is unity for the instructed behaviour, and zero for all others), and the 'decision' is a simple one-to-one mapping between this representation and choice. Determining the specific contribution of these different signals will require more sophisticated tasks that involve nuanced, yet quantifiable, value judgements.

In a 'free choice' task design (FIG. 3c) that is relatively new to the field of neurophysiology, monkeys use eye movements to choose between a pair of visual targets in the absence of any direct instruction about which target should be selected. Instead, such tasks place animals in a situation similar to that faced by our fisherman, in which the valuation and decision transformations of FIG. 2b are the primary determinants of behaviour. Paradoxically, the first challenge in designing such tasks is to create a situation in which behaviour is not actually 'free', at least in the sense of being uncontrolled. For these studies to be informative, behaviour must be predictable on the basis of each task's reward structure, making it possible to quantify the underlying value transformation.

Demonstrating behavioural control. The logic here is borrowed from sensory psychophysics, in which an animal's perceptual sensitivity is captured by a psychometric function that relates behavioural choices to the strength of a sensory signal (see, for example, FIG. 3b). Well-behaved psychometric functions provide confidence that behaviour is under 'stimulus control', allowing the exploration of underlying neural signals as the animal performs the task. By analogy, before embarking on a search for the neural correlates of the valuation and decision transformations indicated in FIG. 2b, we must first create a situation in which behaviour is lawfully related to some experimentally manipulated determinant of value, allowing us to monitor an animal's covert valuation through its overt behaviour.

Three recent studies[76–78] take two different approaches to effecting reliable value-based behaviour, one drawn from the theory of competitive games, the other from a general principle of animal foraging behaviour. Both approaches allow us to predict the average frequencies with which an animal should choose response alternatives if it is sensitive to the statistics that govern reward availability in its environment and behaves so as to maximize its overall rate of reward. In the setting of competitive games[76,78], these predictions derive from Nash equilibria, whereas in the foraging context[77], they emerge directly from Herrnstein's matching law (BOX 3). In either case, the correspondence between predicted and actual behaviour provides a means of assessing behavioural control (FIG. 3d).

Understanding local strategy. Behavioural control is necessary but not sufficient to formulate meaningful questions or conclusions about neural activity. This should come as no surprise: the predictions of the Nash equations or the matching law pertain to average behaviour at equilibrium. By contrast, the goal of neuroscience is to determine the neural mechanisms responsible for behavioural dynamics — the local behavioural strategies that give rise to average behaviours by dictating an animal's individual choices. This distinction between molar (average) and molecular (local) level descriptions of behaviour poses a major problem because, in the case of both the matching law and the Nash equilibrium, the local behavioural strategies that produce these average behavioural phenomena remain

© 2005 Nature Publishing Group

Box 2 | **A tale of two systems?**

Many accounts of more complex behaviour invoke the existence of two distinct decision-making systems, each having a dominant role in valuation and action selection under different circumstances. In behavioural economics, Kahneman and Tversky distinguish between systems that are responsible for intuitive and deliberative judgements[93]. In the field of reinforcement learning, an analogous distinction is made between systems that support the learning of stimulus–response bonds (habits), which correspond to the intuitive system of Kahneman and Tversky, and those that participate in goal-directed action (instrumental control), which correspond to the deliberative system[94,95]. The intuitive/habit system is seen to be responsible for simple behavioural routines that are learned slowly after repeated experience, possibly through the types of dopamine-mediated associative mechanism that are outlined in the main text. The system responsible for goal-directed action, meanwhile, is proposed to rationally evaluate alternative action–outcome scenarios, to rapidly adapt to changing reinforcement contingencies and to be mediated by distributed (and largely unspecified) circuits within prefrontal and association cortices.

As an explanatory heuristic, this dichotomy has some appeal, and is consistent with the divergent effects of certain experimental manipulations on instrumental versus habitual behaviours[96,97]. However, at the physiological level, there is little direct evidence for such a strict division. Midbrain dopamine neurons broadcast their error signal to both the striatum and the cortex, and as the functional scope of that signal has yet to be defined, biologically it seems more appropriate to view such decision-making systems not as distinct entities, but as the theoretical extremes of a continuum of processes that are responsible for valuation and action.

poorly characterized. Therefore, any systematic approach to studying valuation must begin by extracting from the behaviour itself a quantitative model of the local algorithm used by the animal. Ideally, this model should reveal the 'hidden variables' that link reward history to behaviour. These are the variables that govern the animal's choice on each trial, and the goal of carrying out neurophysiological experiments should be to discover how and where these variables are computed and represented at the neural level. We now examine three free-choice studies from the perspective of these three key elements — behavioural control, the modelling of local choice algorithms and physiology.

Value signals in frontal cortex. Barraclough and colleagues[76] investigated neural activity in the context of a version of the 'matching pennies' game. In every round, each of two players places a coin on the table. When the coins are revealed, both go to player A if their faces match and to player B if their faces differ. FIGURE 4a shows how this game is described in terms of a payoff matrix — a standard way to represent the structure of a competitive game. In the experiments, a monkey makes an eye movement to choose one of two targets while the computer also selects a target. The monkey receives a drop of juice only if these selections match. The Nash equilibrium strategy in this task is for each player to choose each of the two targets randomly and with equal probability.

Barraclough and colleagues provide a particularly clear demonstration of the distinction between average behaviour and the local choice algorithm used at the level of individual trials. In separate experiments, the computer opponent used two different algorithms for its choice of targets, leading the monkeys to adopt very different choice strategies. In the first algorithm, the computer acted naively, predicting the monkey's next choice by evaluating only the animal's recent choices without regard for rewards. In response, the monkey adopted a simple strategy in which it chose the same target if the previous choice was rewarded but switched to the other target if the previous choice was unrewarded (a 'win-stay-lose-switch' strategy). In the more sophisticated algorithm, the computer considered the monkey's rewards on previous trials as well as its choices. This completely negated the win-stay-lose-switch strategy and caused the animals to choose in a manner that is characteristic of equilibrium mixed strategies (BOX 3), in which choice seems to be governed by the unpredictable toss of a weighted coin. Importantly, despite these marked differences in local choice strategy, in both versions of the task the monkeys' average frequency of choosing the two targets was 50/50 and therefore consistent with the Nash equilibrium. This indicates that analysis limited to the level of equilibrium behaviour is unlikely to be sufficient to support the meaningful interpretation of the neurophysiological signals that underlie choice behaviour.

Analysing the version of the task in which the animals' choices do seem random in key respects, Barraclough and colleagues nevertheless found that they could predict the monkeys' local probability of choosing one or other target using a reinforcement learning algorithm[38,79]. Consistent with the general framework in FIG. 2b, this model implements a value transformation by maintaining an ongoing representation of the recent value of each target, which it updates on the basis of the monkeys' choice and outcome on the previous trial. The difference in these value representations provides an ongoing estimate of an animal's probability of choosing each alternative. On each trial, this estimate provides the input to a probabilistic decision rule that dictates the ultimate choice that the animal makes. In essence, the relative values of the two targets, computed from the local history of choices and rewards, capture the 'weight' of a coin being tossed to render individual decisions (BOX 4).

This model of local choice provided the authors with a framework for interpreting neural signals recorded from DLPFC. These recordings revealed subpopulations of cells, the activity of which reflected the monkey's choice on the preceding trial, its outcome (reward or not) and the conjunction between the two. The last of these signals is exactly the input needed to update the target-specific value representations predicted in the authors' reinforcement learning model. Taken as a whole, these neurophysiological data indicate that DLPFC might contribute appropriate input to the value tranformation operating in this task, but that the value representations themselves are likely to be computed and stored in downstream areas. Nevertheless, this study shows nicely how an approach that incorporates the three key elements — value-based behavioural control, quantitative modelling of local choice strategy and the neurophysiological exploration of model variables — can be fruitfully

© 2005 Nature Publishing Group

Box 3 | **Assessing covert valuation through overt behaviour**

Behaviour in competitive games and Nash equilibria

As in the schoolyard game of 'rock–paper–scissors', outcomes in strategic games depend not only on an individual's own actions, but also on those of their opponents. The mathematician John Nash showed that in such situations, there is at least one set of strategies among players that is stable, in the sense that no individual player can do better by deviating from their chosen strategy[98]. Behaviour at such 'Nash equilibria' frequently comprises what is termed a 'mixed' strategy, meaning that players' choices can be described by a probability density function over available options. Mixed strategies have two interesting characteristics: they equalize payoffs across alternatives (each player experiences a constant average return for their choices), and they yield behaviour that is stochastic, or unpredictable in time, as though sequential choices were determined by a homogeneous Poisson process (for example, a flip of a weighted coin). This unpredictability is driven by the competition between players, as any regularity in one player's behaviour might be detected and exploited by another. Originally conceived as a prescriptive theory of strategic behaviour, Nash's equations were subsequently found to accurately describe actual behaviour in many real world competitive situations[99]. Both Barraclough and colleagues[76] and Dorris and Glimcher[78] have exploited this fact to generate quantitative predictions about the behaviour of monkeys engaged in strategic conflict against a computer opponent.

Animal foraging behaviour and Herrnstein's matching law

In our own experiments[77], we have used a principle of animal foraging behaviour known as the 'matching law'[100,101], which predicts that animals allocate their time or responses in direct proportion to the fraction of total rewards earned from alternative options. Matching behaviour has been most thoroughly documented in the context of concurrent variable interval reward schedules, in which alternatives are baited with rewards probabilistically and only become un-baited when the animal selects the alternative and collects the reward. In such situations, matching approximates the optimal behavioural strategy[102,103]. Furthermore, like mixed strategies at the Nash equilibrium, matching is characterized by behaviour that is stochastic[103,104], and that equalizes the average payoff or return experienced for choices of each alternative. Like Nash equilibria, the matching law is useful because it allows us to construct quantitative expectations about behaviour in the context of foraging games.

applied to the problem of uncovering value-related neural signals.

Value signals in parietal cortex. The remaining two studies focus on value representations in area LIP. Discussed earlier in the context of perceptual discrimination tasks, LIP has strong connections to DLPFC and shows persistent activity similar to that found in prefrontal cortex. Individual LIP neurons become active in advance of shifts in an animal's gaze[67–69,80] or focus of attention[69,81,82] to a particular region of space, termed the neuron's 'response field'. To study the involvement of LIP in free-choice tasks, on each trial one of the response targets is placed within the response field of the cell under investigation; a given behavioural choice therefore results in an eye movement either towards or away from the cell's response field. At issue here is not whether the cell will encode this planned eye movement — we know that to be the case — but whether its activity will also convey information about the valuation process that lies upstream of that plan.

As discussed earlier, LIP firing rates are modulated by many factors, leading to the proposal that individual LIP neurons integrate information from multiple sources to encode their combined influence on the likelihood of an eye movement to the cell's response field. This perspective is reminiscent of the concept of a 'common currency', which was raised earlier in the context of convergent inputs to the reward system. By analogy, we might conceive of LIP activity as a common currency for the planning of orienting responses, mapping input from multiple sources that might contribute to the selection of upcoming eye movements onto a single common representation (BOX 1). Therefore, value-related signals should be evident in LIP for any task that requires eye movements to be selected according to the value of competing targets. Importantly, these value signals should reflect the specifics of the underlying value transformation, and provide a window onto the otherwise hidden decision variables that link reward experience to action.

The perspective of behavioural dynamics. In our own experiments[77], we record from LIP in a task in which average behaviour conforms to the predictions of the matching law (BOX 3). In this task, monkeys use eye movements to choose between competing alternatives that are rewarded with probabilities that change unpredictably across blocks of trials. The monkeys are sensitive to these shifts in reward probability and adjust the fraction of responses that they make to each alternative to 'match' the fraction of rewards experienced from that option on the current block (FIG. 5a). Like Barraclough and colleagues, we find that a simple strategy based on recent reward experience describes the local mechanism that is responsible for animals' choices in this task.

This strategy takes the form of a linear–nonlinear probabilistic (LNP) model (FIG. 6a). At its input stage, the model uses leaky integrators of reward to compute the local income due to each response alternative. The model's subsequent nonlinear stage uses simple division to normalize these estimates of local income. The resulting variable — 'local fractional income' — captures the animal's relative valuation of each response alternative, and directly dictates its instantaneous probability of choosing that option on the current trial. Like the model of Barraclough and colleagues, our LNP model renders a final choice on each trial through the toss of a biased coin weighted with this probability (BOX 4). This model is similar to Herrnstein's original matching law in that it relates probability of choice to the fraction of rewards due to each alternative. However, as we emphasized earlier, there is an important difference between the levels at which these two descriptions are intended to capture the underlying behaviour. The classic matching law is intended as a molar or average description of equilibrium behaviour, and is used to predict the steady state that will emerge after an animal has been exposed to a particular set of reward contingencies for a long period of time. By contrast, the local model detailed above is intended as a molecular or mechanistic description of behavioural dynamics that can be used to estimate local value and to predict the animal's individual

© 2005 Nature Publishing Group

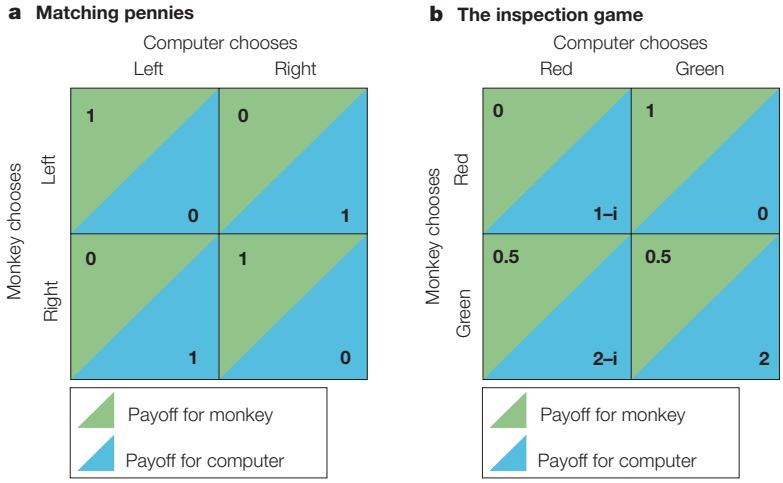

Figure 4 | **Payoff matrices for competitive games.** For a trial in the matching pennies game (**a**) or the inspection game (**b**), a payoff matrix defines the outcome for each player on the basis of the combined actions of both players. Green and blue represent the payoff experienced by the monkey and computer, respectively, for each possible combination of choices. In the inspection game, 'i' defines a cost to the computer for choosing ('inspecting') the red ('risky') target. By manipulating this cost across blocks, the mixed strategy predicted by the Nash equilibrium can be changed. Panel **a** adapted, with permission, from REF. 76 © (2004) Macmillan Publishers Ltd. Panel **b** adapted, with permission, from REF. 78 © (2004) Elsevier Science.

choices both at steady state and during times of behavioural change.

The key parameter in this model is the time constant of reward integration, which captures the decaying influence of past rewards on current choice. In simulations, the percentage of potential rewards collected by this model critically depends on this parameter. Interestingly, we find that the specific integration time that the monkeys deploy is approximately optimal given the statistics that govern the availability of rewards in the task. Using the monkeys' integration time, the model accurately predicts the animals' actual choice behaviour (FIG. 6b), and can itself generate behaviour that is qualitatively similar to that of the animals (FIG. 6c). Most importantly, this model provides us with candidate decision variables for physiological investigation. On each trial, it captures the animal's strategy and distills its relative valuation of the response alternatives into a single number — local fractional income. It is this valuation variable that strongly modulates LIP firing rates.

FIGURE 6d illustrates this result for a single representative LIP neuron recorded during performance of the matching task. Each point represents the cell's mean firing rate on a single trial, plotted against the local fractional income of the colour located within the cell's response field. Blue and green points identify trials on which the monkey chose the target located inside or outside the cell's response field, respectively. The cell reliably predicts the monkey's eye movement responses, firing more strongly before choices of the response field target. However, among trials that share a common final eye movement, the cell is also strongly modulated by fractional income, firing more strongly when the fractional income of the response field target is high, whether or not that target is ultimately chosen. This effect of fractional income is independent of any trial-to-trial variation in the fine details of the monkey's eye movements. Importantly, these signals are apparent to us only because our behavioural model affords access to the animal's underlying value transformation, which is local in time. Our previous attempts to detect value signals in LIP on the basis of global behavioural changes between blocks were marginally successful at best. Like Barraclough and colleagues, we owe our progress to an approach that combined valuation-based behavioural control, modelling of the proximal algorithm that generates individual choices and the neurophysiological study of the variables revealed by this model.

The perspective of behavioural equilibria. Dorris and Glimcher[78] base their experiments on a well-characterized competitive interaction known as 'the inspection game'. The general structure of this task is similar to that used by Barraclough and colleagues: on every trial, the monkey and the computer each select one of two eye movement targets, and the outcome depends on their combined choices. However, in this task the payoff matrix that defines the relationship between choices and outcomes is more complex (FIG. 4b). The monkey faces a choice between a green 'certain' target that delivers a small volume of juice with 100% probability, and a red 'risky' target that sometimes delivers a large volume of juice but at other times delivers nothing. The actual outcome of a 'risky' choice depends on the computer's actions, as the monkey gets nothing on trials in which the computer also selects ('inspects') the risky target. Therefore, the computer can block the monkey from receiving rewards for risky choices, although a 'cost' charged to the computer discourages it from inspecting on every trial. The elegance of this design comes from the fact that by simply manipulating the computer's 'inspection cost', the investigators can change the Nash equilibrium of the game, and, therefore, the proportion of trials on which the monkey is predicted to choose the risky target.

There are clear parallels between this study and the last. As in the matching task, monkeys playing the inspection game react appropriately to the blockwise changes in predicted response probability, adjusting the frequency with which they select the risky target to approximate each new equilibrium point (FIG. 5b). The authors then record LIP activity during task performance, always placing the red 'risky' target within the response field of the cell under investigation. Despite these similarities, the results of the two studies seem to differ, leading the authors to different conclusions about the role of LIP in value-based choice.

Dorris and Glimcher's results are remarkable owing to the absence of an effect on LIP activity. Specifically, whereas the frequency with which their monkeys select the risky (response field) target changes markedly across blocks of trials, LIP firing rates change little, remaining fixed at a fairly constant level whenever the monkey chooses the response field target (FIG. 7c). To interpret

© 2005 Nature Publishing Group

Box 4 | **Probabilistic choice: does the brain play dice?**

Both of the decision pathways shown in FIG. 2 make reference to an explicit representation of 'probability of choice.' It is a matter of some contention whether the brain ever works by computing the probability of alternative actions and subsequently selecting among them at random — as if rendering individual choices through the toss of a weighted coin. Certainly in the realm of perceptual decisions, there is no need to posit explicit probabilism in the decision-making process. Although traditional psychometric functions (FIG. 3b) show a smooth relationship between stimulus parameters and probability of choice, signal detection theory[105] tells us that this apparent probabilism can be explained entirely by noise at the level of the sensors. Even if the underlying decision criterion is perfectly deterministic, noise in the sensory representation of identical stimuli can cause some observations to lie on different sides of the decision criterion, giving rise to unpredictability at the level of observed behaviour. However, although we do not need explicit probabilism on the part of the brain to explain data from studies of sensory-based decision making, neither do these data preclude it. For example, some of the apparent variability in choice, which, according to signal detection theory, is attributable to noise in the sensory representation, might actually arise from true blurring in the internal decision criterion.

Value-based decisions offer a stronger rationale for probabilistic models of choice behaviour. It is often argued from the perspective of game theory that explicitly probabilistic behaviour is required of animals in competitive environments (BOX 3). However, even in many non-competitive foraging situations, like our own matching experiments, the most successful models of behaviour frequently incorporate a probabilistic stage (see FIG. 6). Nevertheless, just as the argument above does not bar true probabilism from sensory decision making, neither do these arguments necessitate its presence in value-based decision making. There might be some better model of choice in our matching task that is actually deterministic, and however optimal it might be to behave randomly in competitive settings, animals might not actually do so.

Future experiments might address how and where the apparent unpredictability in animal behaviour arises. All variability in animal choice might, ultimately, be traced back to noise at the level of sensory and value representations. Alternatively, noise at the level of the decision mechanism might introduce inherent unpredictability. A third possibility is that the brain might explicitly compute the probability of taking each of several actions and possess mechanisms for choosing among them, much like flipping a biased coin. The framework of FIG. 2 is meant only to be general enough to encompass all three of these possible realities.

this result, the authors appeal to the logic of behaviour at the Nash equilibrium, at which point a player's overall distribution of choices should equalize the average payoff (rewards per choice) resulting from alternative actions (BOX 3). On the basis of this assumed equality in payoffs, the authors suggest that the 'subjective desirability' of the two alternatives remains equal across blocks. From this perspective, invariance in LIP firing rates across blocks can be interpreted as evidence that LIP encodes each alternative's average payoff, or subjective desirability (which is a constant), rather than the probability of choosing that alternative (which varies). Although unorthodox in its appeal to the absence of a neural effect as evidence of a correlation, this conclusion does seem to account for an otherwise puzzling pattern of physiological results.

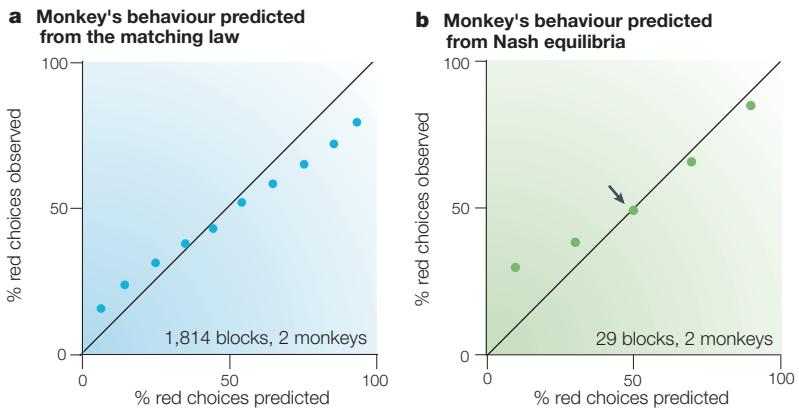

Figure 5 | **Demonstrating behavioural control.** Average monkey behaviour in the matching task (**a**) and the inspection game (**b**) corresponds to the predictions of the matching law and the Nash equilibrium, respectively. In both panels, each point represents an average over blocks of trials with the specified predicted choice probability. Standard error bars are smaller than the data points and are, therefore, not visible. The arrow indicates the point that corresponds to the single mixed strategy explored in the Barraclough study[76]. Panel **a** adapted, with permission, from REF. 77 © (2004) American Association for the Advancement of Science. Panel **b** adapted, with permission, from REF. 78 © (2004) Elsevier Science.

In contrast to Dorris and Glimcher's conclusions, we proposed earlier that LIP activity comprises a common currency for planning shifts in gaze or visual attention. From this perspective, the activity of a cell in area LIP signals the evolving probability or likelihood that the animal will select the target located within that cell's response field. Our data from the matching task are consistent with this interpretation, locating LIP at, or downstream from, the decision transformation that maps value to the probability of a particular choice or motor response (FIG. 2b). By associating LIP activity with subjective desirability rather than the local probability of an eye movement, Dorris and Glimcher's results indicate a very different interpretation. In this view, LIP encodes an abstract representation of the relative desirability of external stimuli apart from any specific motor plan[83,84], implying that LIP is located upstream of the putative decision transformation that links valuation to action (FIG. 2b).

The discovery of a 'pure' representation of value or desirability would comprise an important advance, but we are not convinced that the data support this conclusion. Rather, we believe that the apparent discrepancy results from a misplaced emphasis in the study by Dorris and Glimcher on the global Nash equilibrium rather than on the local process through which that equilibrium is generated. The Nash equilibrium is not

© 2005 Nature Publishing Group

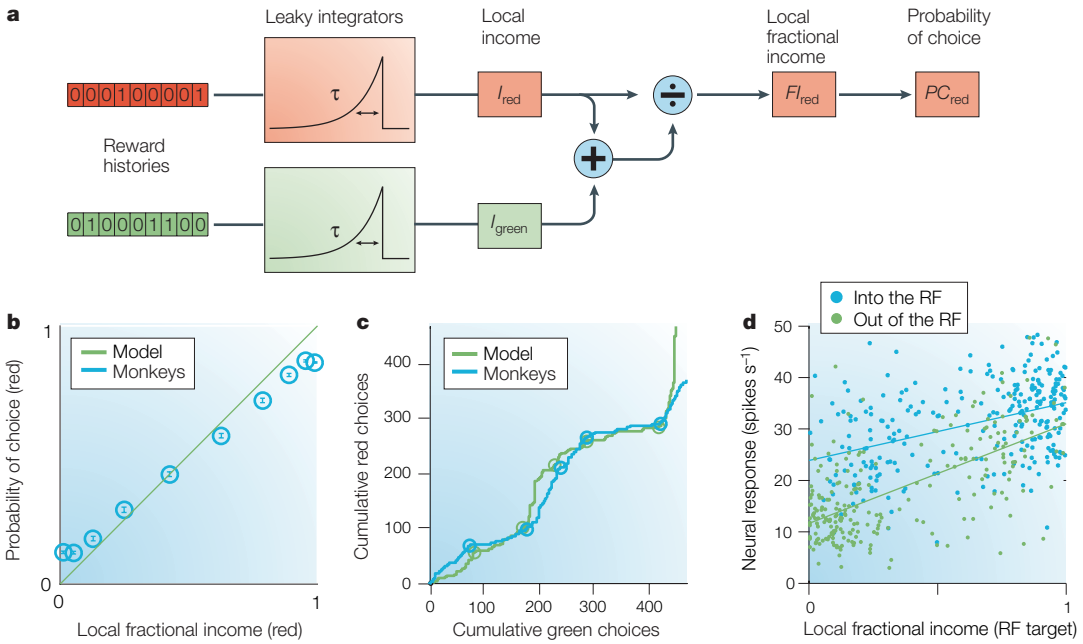

Figure 6 | **A local model of matching behaviour. a** | A linear–nonlinear probabilistic model uses leaky integration over recent reward experience to estimate the local income due to each response option (I_{red}, I_{green}). In a local formulation of Herrnstein's matching law, these estimates are used to compute the local fractional income of each option (that is, FI_{red}), which directly translates into the animal's instantaneous probability of choosing that option on the current trial (that is, PC_{red}). **b** | Predictive sufficiency: model output is predictive of animal choice. **c** | Generative sufficiency: when exposed to an identical block sequence in simulation, model behaviour closely resembles that of the monkey. **d** | The activity of an example lateral intraparietal neuron shows dependence of neural response on the local fractional income of the reponse field (RF) target. Each point represents the cell's mean firing rate on a single trial. Blue and green indicate trials ending with choices into and out of the response field, respectively. Lines are least squares regressions fit to the corresponding points (p for each <0.001). Panels **a** and **d** adapted, with permission, from REF. 77 © (2004) American Association for the Advancement of Science. Panels **b** and **c** reproduced, with permission, from REF. 77 © (2004) American Association for the Advancement of Science.

a mechanism; it is a description of a behavioural end-state measured by averaging choices over time. In any competitive game, individual choices are presumed to result from an underlying dynamic process through which each player continuously tries to gain an advantage over the other. From our perspective, the subject of real interest — and the computation that is likely to have a neural implementation — is the mechanism that generates these individual choices. It is not obvious why the brain should encode a behavioural end-state abstractly. By analogy, a thermostat might regulate the average temperature in a room, but it does so by computing and responding to local temperature fluctuations, not to the average temperature.

Dorris and Glimcher acknowledge that behaviour at the Nash equilibrium is likely to be influenced by a local process of valuation. Rather than develop a model of local valuation, however, they use the computer's inspection algorithm as a proxy for the animal's estimation of the value of the two targets on each trial. This algorithm does not directly incorporate information about the animal's history of rewards, so it offers, at best, a rough approximation of the animal's actual value transformation. Tellingly, LIP activity covaries with local fluctuations in this rough estimate of value (for further information, see REF. 78), and therefore presumably with local fluctuations in probability of choice. We suspect that this result points to the real signal in these data, and that an appropriate model and analysis would reveal a systematic relationship between this signal and the actual local value variable that dictates the animal's instantaneous probability of choice.

Local or global valuation? If we are correct that area LIP represents local value and local probability of choice, an important question remains: why does LIP activity in the inspection game seem to remain roughly constant across blocks of trials despite large variations in the monkey's average probability of choosing the response field target (FIG. 7c)? To explore this issue, we re-examine our own experiments.

On each trial in the matching task, the response field target has both a global value (determined by its global fractional income and computed using blockwise reward rates) and a local value (determined by its local fractional income and computed using our behavioural model; FIG. 6a). By simultaneously sorting trials according to both of these measures of value, we can directly examine the influence of each while controlling for the effect of the other. In this manner, FIG. 7a illustrates the joint influence of global and local value on the average normalized activity of 43 LIP neurons recorded in the matching task. The individual pixels in this plot show average LIP activity for trials on which the response field target had the specified combination of global and local values; contours of constant colour reflect trials on

© 2005 Nature Publishing Group

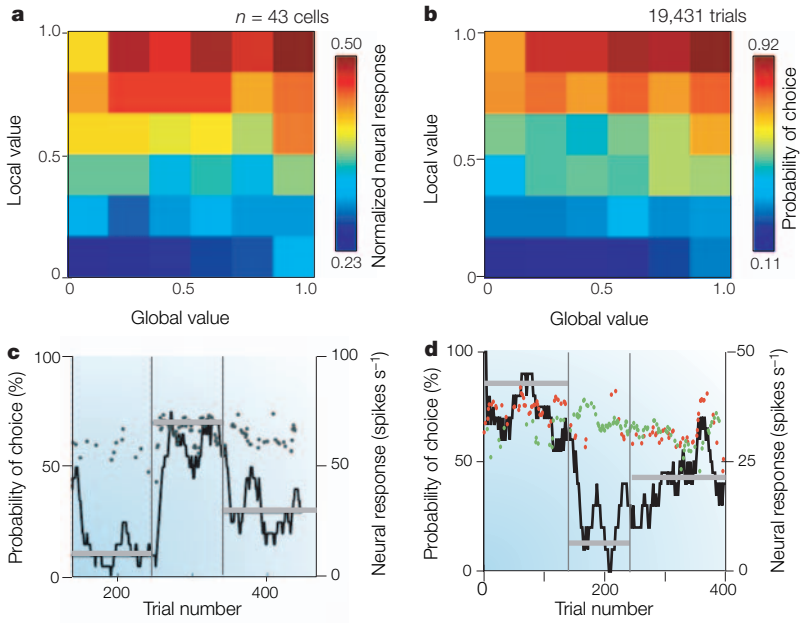

Figure 7 | **Influence of global and local values on monkey choices and lateral intraparietal area activity. a** | Average normalized neural response of 43 lateral intraparietal (LIP) neurons in the matching task as a function of both the global (abscissa) and local (ordinate) values of the response field target. **b** | Monkey's probability of choosing the response field target in the matching task as a function of both the global (abscissa) and local (ordinate) values of the response field target. **c**,**d** | Activity of an example LIP neuron during performance of three blocks of the inspection game (**c**) or the matching task (**d**). For both panels, thin vertical black lines represent unsignalled block transitions; horizontal lines indicate the average percentage of red choices predicted by either the Nash equilibrium (**c**) or the matching law (**d**); black jagged lines represent a 20-trial running average of the monkey's probability of choosing the red target; dots indicate a 20-trial running average of neural activity that includes only those trials on which the monkey chose the response field target. In **d**, dot colour indicates the colour of the chosen target; in **c**, the red 'risky' target was always positioned within the response field. At first glance, the result in **b** might seem inconsistent with FIG. 5a, which showed a linear relationship between probability of choice and global value. The key to resolving this apparent discrepancy is that trials from any given global value group are distributed non-uniformly across the range of local values. Therefore, most of the trials that contribute to the rightmost point in FIG. 5a (the highest global value) come from the upper right-hand side of the figure shown here — trials that also have high local value. Similarly, most of the trials that contribute to the leftmost point in FIG. 5a (the lowest global value) come from the lower left-hand side of the figure shown here — trials that also have low local value. Bins in the other two quadrants are more sparsely populated. Therefore, the relationship between choice and global value in FIG. 5a is entirely due to a relationship between choice and local value and the presence of progressively more trials with high local values as we select for greater global value. Panel **c** adapted, with permission, from REF. 78 © (2004) Elsevier Science. Data in panels **a**, **b** and **d** from REF. 77.

which neural activity was also constant. The horizontal orientation of these contours shows that activity in LIP is strongly influenced by local value but essentially independent of global value (note that this same result is obtained if trials that end in responses directed either towards or away from the response field target are analysed separately). FIGURE 7b reveals a similar pattern of effects for the monkey's probability of choice: the monkey chooses the response field target when that target has a high local value, and is only weakly influenced by the target's global value.

We believe that this inherently local valuation process, when combined with bias in the distribution of trials sampled, can account for the apparent discrepancy between our LIP results and those described by Dorris and Glimcher. Like idealized behaviour at the Nash equilibrium, perfect matching behaviour is characterized by equivalence in the average payoff or return experienced from competing response alternatives (BOX 3). Following the example of Dorris and Glimcher, therefore, we might interpret the fact that LIP firing rates do not vary with changes in global value (FIG. 7a) as evidence that LIP activity allows the equal subjective desirability of the response targets to be encoded. However, we believe that the actual explanation for this result is more mundane: changes in global value seem to exert little effect because the influence of global value on both choice and neural activity is mediated entirely through effects on local value. To reiterate, when the global value of the response field target changes across blocks, it changes the frequency with which the response field target has a high local value, but the relationship between local value and choice, and consequently the distribution of local values on trials when the monkey actually chooses the response field target, changes little, if at all.

This result means that any analysis that is restricted only to trials on which the monkey chooses the response field target will be biased towards trials on which the local value of the response field target is high, and will consequently give the appearance of constant (and relatively high) LIP firing rates. This effect is illustrated directly in FIG. 7d, where we re-plot the data from the same matching experiment presented in FIG. 6d in the format used by Dorris and Glimcher to present their inspection game data (reproduced in FIG. 7c). Each dot in this plot represents a running average of the cell's activity, but only considers trials on which the monkey actually chose the response field target. Like the corresponding figure in Dorris and Glimcher's paper, the resulting plot is remarkable for the constant high level of neural activity. We have already seen, however, that these same data contain real variance that is well explained by a local mechanism (FIG. 6d) that makes no appeal to explicit neural encoding of the subjective equivalence implied by the equilibrium state. The constant firing rates that we observe in FIG. 7d are a direct consequence of the simple fact that both neural activity and behaviour track changes in local value. In the absence of evidence to the contrary, we believe that LIP data obtained in the context of the inspection game can also be explained by local value computations that directly dictate the animal's probability of choosing one or other target on any given trial (FIG. 2b).

The inspection game incorporates both the dynamics of competition and the challenge of an environment that changes unpredictably over time. As such, it is likely to elicit a local strategy that differs from that seen in either of the previous studies. Uncovering the details of this strategy and its accompanying value transformation will be an important challenge for future work. This knowledge might demystify some of the hidden behavioural processes that govern competitive interactions and provide a powerful tool for exploring their underlying neural substrates.

© 2005 Nature Publishing Group

Concluding remarks

Although Jean-Dominique Bauby was capable of executing just a single meaningful action — an eyeblink — he nevertheless describes his inner mental life in terms of a rich behavioural repertoire. His description suggests a fundamental inability to separate abstract decision making from the concrete push and pull of motor planning. In a similar vein, it has been suggested that the existence of decision-related signals in area LIP reflects a neural architecture in which decisions are computed and represented within the very neural structures that guide behavioural responses[10,85–89]. This idea indicates that decisions and the information that contributes to them are ultimately expressed in the currency of potential or deferred actions. In this regard, it is certainly provocative that decision-related activity has been explored most profitably in regions of the brain that are responsible for the planning of particular motor responses. However, futher experiments will be needed to reveal whether decision variables such as those related to valuation also have more abstract response-invariant representations in the brain.

We have argued that the key to understanding the neural basis of cognitive phenomena — like valuation and choice — is to move putative internal variables such as 'value' out of the realm of intuition and into the realm of quantitative measurement and prediction. At a practical level, this necessitates experimental approaches that place an animal's behaviour under the verifiable control of the variable of interest. In the context of value-based choice, we have discussed how global behavioural phenomena such as Nash equilibria or matching behaviour can provide a means to this end. We have emphasized, however, the crucial importance of understanding the proximal decision mechanisms that give rise to these global phenomena, and the central role of quantitative models of choice behaviour in generating this understanding. These models are our means of accessing the internal decision variables that govern local choice, and, in the final analysis, it is these variables that should be the focus of neurophysiological and imaging investigations.

At several places in this review, we draw parallels between perceptual and value-based choice. More than a convenient organizational tool, this parallelism reflects our belief that new efforts to understand value-based decision making might bring together two areas of neuroscience that have traditionally existed in separate spheres — the study of perception and cognition, and the study of reward and motivation. This partition is largely an artefact of history, reflecting separate research agendas born of pioneering discoveries in systems neuroscience. Neurobiological systems themselves do not recognize these borders, and ultimately need to be studied and understood from a single coherent perspective. With its inherent emphasis on the intersection of cognitive and motivational processes, the study of value-based choice might be uniquely positioned to lay the foundations for this unified neurobiology of choice behaviour.

1. Bauby, J. D. *The Diving Bell and The Butterfly: A Memoir of Life in Death* (Alfred A. Knopf Inc., New York, 1997).
2. Glimcher, P. W. & Rustichini, A. R. Neuroeconomics: the consilience of brain and decision. *Science* **306**, 447–452 (2004).
3. Parker, A. J. & Newsome, W. T. Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* **21**, 227–277 (1998).
An exploration of the link between neural activity and sensory perception, summarizing work that comprises the background to recent neurophysiological studies of perceptual decision making.
4. Rizzolatti, G. & Luppino, G. The cortical motor system. *Neuron* **31**, 889–901 (2001).
5. Graziano, M. S., Taylor, C. S., Moore, T. & Cooke, D. F. The cortical control of movement revisited. *Neuron* **36**, 349–362 (2002).
6. Andersen, R. A. & Bruneo, C. A. Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* **25**, 189–220 (2002).
7. Leon, M. I. & Shadlen, M. N. Exploring the neurophysiology of decisions. *Neuron* **21**, 669–672 (1998).
8. Georgopoulos, A. P. Neural aspects of cognitive motor control. *Curr. Opin. Neurobiol.* **10**, 238 (2000).
9. Schall, J. D. Neural basis of deciding, choosing and acting. *Nature Rev. Neurosci.* **2**, 33–42 (2001).
10. Gold, J. I. & Shadlen, M. N. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16 (2001).
Proposes that neural systems implement categorical decisions about perceptual stimuli by computing a decision variable related to the logarithm of the likelihood ratio in favour of one or another alternative. Discusses how this framework can be used to interpret neural signals recorded from eye movement planning centres during perceptual discrimination tasks.
11. Glimcher, P. W. Making choices: the neurophysiology of visual-saccadic decision making. *Trends Neurosci.* **24**, 654–659 (2001).
12. Romo, R. & Salinas, E. Touch and go: decision making mechanisms in somatosensation. *Annu. Rev. Neurosci.* **24**, 107–137 (2001).
13. Romo, R., Hernandez, A., Zainos, A., Lemus, L. & Brody, C. Neural correlates of decision-making in secondary somatosensory cortex. *Nature Neurosci.* **5**, 1217–1225 (2002).
14. Romo, R. & Salinas, E. Flutter discrimination: neural codes, perception, memory and decision making. *Nature Rev. Neurosci.* **4**, 203–218 (2003).
15. Graham, N. V. S. *Visual Pattern Analysers* (Oxford Univ. Press, Oxford, 1989).
16. Shadlen, M. N. & Newsome, W. T. Motion perception: seeing and deciding. *Proc. Natl Acad. Sci. USA* **93**, 628–633 (1996).
One of the first demonstrations of a correlation between the activity of single cortical neurons and an evolving perceptual decision.
17. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
18. Kim, J. N. & Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neurosci.* **2**, 176–185 (1999).
19. Horwitz, G. D. & Newsome, W. T. Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *J. Neurophysiol.* **86**, 2543–2558 (2001).
20. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).
21. Hanks, T. D. & Shadlen, M. N. Microstimulation of macaque area LIP affects decision making in a motion discrimination task. *Soc. Neurosci. Abstr.* 20.9 (2004).
22. Bisley, J. W. & Goldberg, M. E. Neuronal activity in the lateral intraparietal area and spatial attention. *Science* **299**, 81–86 (2003).
23. Platt, M. L. & Glimcher, P. W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1998).
24. Kahneman, D. & Tversky, A. (eds) *Choices, Values, and Frames* (Cambridge Univ. Press, Cambridge, 2000).
25 Berridge, K. C. & Robinson, T. E. Parsing reward. *Trends Neurosci.* **26**, 507–513 (2003).
Discusses the distinct psychological components of reward and their underlying neural substrates. Emphasizes the distinction between systems devoted to the affective ('liking') and motivational ('wanting') aspects of reward, and between motivational processes based on simple Pavlovian associations and those involving more cognitive representations of value.
26. Olds, J. & Milner, P. M. Positive reinforcement produced by electrical stimulation of septal area and other regions of the rat brain. *J. Comp. Physiol. Psychol.* **47**, 419–427 (1954).
27. Shizgal, P. Neural basis of utility estimation. *Curr. Opin. Neurobiol.* **7**, 198–208 (1997).
Summarizes findings from a series of elegant experiments involving BSR in the rat that indicate the existence of a final common neural representation for reward information.
28. Wise, R. A. & Bozarth, M. A. A psychomotor stimulant theory of addiction. *Psychol. Rev.* **94**, 469–492 (1987).
29. Hyman, S. E. & Malenka, R. C. Addiction and the brain: the neurobiology of compulsion and its persistence. *Nature Rev. Neurosci.* **2**, 695–703 (2001).
30. Kelley, A. E. Memory and addiction. Shared neural circuitry and molecular mechanisms. *Neuron* **44**, 161–167 (2004).
31. Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* **28**, 309–369 (1998).
32. Garris, P. A. *et al.* Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature* **398**, 67–69 (1999).
33. Pecina, S., Cagniard, B., Berridge, K. C., Aldridge, J. W. & Zhuang, X. Hyperdopaminergic mutant mice have higher 'wanting' but not 'liking' for sweet rewards. *J. Neurosci.* **23**, 9395–9402 (2003).
34. Liu, Z. *et al.* DNA targeting of rhinal cortex D2 receptor protein reversibly blocks learning of cues that predict reward. *Proc. Natl Acad. Sci. USA* **101**, 12336–12341 (2004).
35. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913 (1993).
36. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).

© 2005 **Nature Publishing Group**

37. Hollerman, J. R. & Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neurosci.* **1**, 304–309 (1998).
38. Sutton, R. S. & Barto, A. G. *Reinforcement Learning* (MIT Press, Cambridge, Massachusetts, 1998).
39. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience* Ch. 9 (MIT Press, Cambridge, Massachusetts, 2001).
40. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
Reviews evidence indicating that midbrain dopamine neurons signal an error in the prediction of future reward (see references 35–37). Proposes that the function of these neurons is particularly well described by a specific class of reinforcement learning algorithms, and shows how a model that uses a dopamine-like signal to implement such an algorithm can learn to predict future rewards and guide action selection.
41. Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241–263 (2002).
42. Montague, P. R., Dayan, P., Person, C. & Sejnowski, T. Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* **377**, 725–728 (1995).
43. Barto, A. G. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 215–232 (MIT Press, Cambridge, Massachusetts, 1995).
44. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neurosci.* **1**, 411–416 (1998).
45. Lauwereyns, J., Watanabe, K., Coe, B. & Hikosaka, O. A neural correlate of response bias in monkey caudate nucleus. *Nature* **418**, 413–417 (2002).
46. Gold, J. I. Linking reward expectation to behavior in the basal ganglia. *Trends Neurosci.* **26**, 12–14 (2003).
47. Montague, P. R. & Berns, G. S. Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284 (2002).
Among the first of a series of recent attempts (see also references 41,48,49) to integrate our emerging understanding of the role of dopamine with other reward-related neural signals for the purpose of developing a coherent framework for understanding the neural basis of reward and valuation.
48. McClure, S. M., Daw, N. D. & Montague, P. R. A computational substrate for incentive salience. *Trends Neurosci.* **26**, 423–428 (2003).
49. Montague, P. R., Hyman, S. E. & Cohen, J. D. Computational roles for dopamine in behavioral control. *Nature* **431**, 760–767 (2004).
50. Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y. & Hikosaka, O. Dopamine neurons can represent context-dependent prediction error. *Neuron* **41**, 269–280 (2004).
51. Rolls, E. T., Critchley, H., Mason, R. & Wakeman, E. A. Orbitofrontal cortex neurons: role in olfactory and visual association learning. *J. Neurophysiol.* **75**, 1970–1978 (1996).
52. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
53. Tremblay, L. & Schultz, W. Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *J. Neurophysiol.* **83**, 1877–1885 (2000).
54. Watanabe, M. Reward expectancy in primate prefrontal neurons. *Nature* **382**, 629–632 (1996).
55. Leon, M. I. & Shadlen, M. N. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* **24**, 415–425 (1999).
56. Kobayashi, S., Lauwereyns, J., Koizumi, M., Sakagami, M. & Hikosaka, O. Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *J. Neurophysiol.* **87**, 1488–1498 (2002).
57. Wallis, J. D. & Miller, E. K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**, 2069–2081 (2003).
58. Roesch, M. R. & Olson, C. R. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* **90**, 1766–1789 (2003).
59. Roesch, M. R. & Olson, C. R. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* **304**, 307–310 (2004).
60. Shidara, M. & Richmond, B. J. Anterior cingulate: single neuron signals related to the degree of reward expectancy. *Science* **296**, 1709–1711 (2002).
61. McCoy, A. N., Crowley, J. C., Haghighian, G., Dean, H. L. & Platt, M. L. Saccade reward signals in posterior cingulate cortex. *Neuron* **40**, 1031–1040 (2003).
62. Ito, S., Stuphorn, V., Brown, J. W. & Schall, J. D. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* **302**, 120–122 (2003).
63. Fuster, J. in *Cerebral Cortex* (eds Peter, J. & Jones, E.) 151–177 (Plenum, New York, 1985).
64. Segraves, M. A. & Goldberg, M. E. Functional properties of corticotectal neurons in the monkey's frontal eye field. *J. Neurophysiol.* **58**, 1387–1419 (1987).
65. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989).
66. Romo, R., Brody, C. D., Hernandez, A. & Lemus, L. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* **339**, 470–473 (1999).
67. Gnadt, J. W. & Andersen, R. A. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* **70**, 216–220 (1988).
68. Barash, S., Bracewell, R. M., Fogassi, L., Gnadt, J. W. & Andersen, R. A. Saccade-related activity in the lateral intraparietal area. I. Temporal properties; comparison with area 7a. *J. Neurophysiol.* **66**, 1095–1108 (1991).
69. Colby, C. L., Duhamel, J. R. & Goldberg, M. E. Visual, presaccadic, and cognitive activation of single neurons in monkey intraparietal area. *J. Neurophysiol.* **76**, 2841–2852 (1996).
70. Bechara, A., Tranel, D., Damasio, H. & Damasio, A. R. Deciding advantageously before knowing the advantageous strategy. *Science* **275**, 1293–1295 (1997).
71. Iversen, S. D. & Mishkin, M. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp. Brain Res.* **11**, 376–386 (1970).
72. Rolls, E. T. The orbitofrontal cortex and reward. *Cereb. Cortex* **10**, 284–294 (2000).
73. Damasio, A. R. *Descartes' Error: Emotion, Reason, and the Human Brain* (Grosset/Putnam, New York, 1994).
74. Musallam, S., Corneil, B. D., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive control signals for neural prosthetics. *Science* **305**, 258–262 (2004).
75. Schultz, W. Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioral ecology. *Curr. Opin. Neurobiol.* **14**, 139–147 (2004).
An up-to-date survey of the types and locations of reward-related signals that have been documented using neurophysiological techniques.
76. Barraclough, D. J., Conroy, M. L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neurosci.* **7**, 404–410 (2004).
77. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787 (2004).
78. Dorris, M. C. & Glimcher, P. W. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* **44**, 365–378 (2004).
References 76–78 describe studies that all used a novel free-choice approach to investigate the behavioural and neural basis of value-based decision making in awake monkeys. These papers are discussed at length in the main text of this review.
79. Erev, I. & Roth, A. E. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* **88**, 848–881 (1998).
80. Platt, M. L. & Glimcher, P. W. Response fields of intraparietal neurons quantified with multiple saccadic targets. *Exp. Brain Res.* **121**, 65–75 (1998).
81. Gottlieb, J. P., Kusunoki, M. & Goldberg, M. E. The representation of visual salience in monkey parietal cortex. *Nature* **391**, 481–484 (1998).
82. Colby, C. L. & Goldberg, M. E. Space and attention in parietal cortex. *Annu. Rev. Neurosci.* **22**, 319–349 (1999).
83. Glimcher, P. W. The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.* **26**, 133–179 (2003).
84. Glimcher, P. W., Dorris, M. C. & Bayer, H. M. Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.* (in the press).
85. Gold, J. I. & Shadlen, M. N. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* **36**, 299–308 (2002).
86. Gold, J. I. & Shadlen, M. N. Representation of a perceptual decision in developing oculomotor commands. *Nature* **404**, 390–394 (2000).
87. Shadlen, M. N. Pursuing commitments. *Nature Neurosci.* **5**, 819–821 (2002).
88. Horwitz, G. D., Batista, A. P. & Newsome, W. T. Representation of an abstract perceptual decision in macaque superior colliculus. *J. Neurophysiol.* **91**, 2281–2296 (2004).
89. Rizzolatti, G. & Craighero, L. The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).
90. Maunsell, J. H. Neuronal representations of cognitive state: reward of attention? *Trends Cogn. Sci.* **8**, 261–265 (2004).
91. Koch, C. & Ullman, S. Shifts in selective visual attention: toward the underlying neural circuitry. *Hum. Neurobiol.* **4**, 219–227 (1985).
92. Itti, L. & Koch, C. Computational modeling of visual attention. *Nature Rev. Neurosci.* **2**, 194–203 (2001).
93. Kahneman, D. in *Les Prix Nobel 2002* (ed. Frangsmyr, T.) 416–499 (2002).
A clear summary of the pioneering work of Kahneman and Tversky on the psychology of valuation and choice, which emphasizes the importance of intuitive over deliberative decision making and challenges basic economic assumptions of rationality.
94. Dayan, P. & Balleine, B. W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298 (2002).
95. Dickinson, A. in *Animal Learning and Cognition* (ed. Mackintosh, N. J.) 45–79 (Academic, Orlando, 1994).
96. Nevin, J. A. Analyzing Thorndike's law of effect: the question of stimulus-response bonds. *J. Exp. Anal. Behav.* **72**, 447–450 (1999).
97. Colwill, R. M. & Rescorla, R. A. Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *J. Exp. Psychology Anim. Behav. Process.* **11**, 520–536 (1985).
98. Nash, J. F. Equilibrium points in n-person games. *Proc. Natl Acad. Sci. USA* **36**, 48–49 (1950).
99. Holt, C. A. & Roth, A. E. The Nash equilibrium: a perspective. *Proc. Natl Acad. Sci. USA* **101**, 3999–4002 (2004).
100. Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4**, 267–272 (1961).
101. Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton Univ. Press, Princeton, New Jersey, 1986).
102. Baum, W. M. Optimization and the matching law as accounts of instrumental behavior. *J. Exp. Anal. Behav.* **36**, 387–401 (1981).
103. Gallistel, C. R. & Gibbon, J. Time, rate and conditioning. *Psychol. Rev.* **107**, 289–344 (2000).
A refreshing, if controversial, look at the subject of operant choice. Argues for distinct decision mechanisms in choice situations that involve opting for the single best alternative, compared with those settings in which an animal must decide in what proportion to allocate its time or responses between alternatives.
104. Heyman, G. M. A Markov model description of changeover probabilities on concurrent variable-interval schedules. *J. Exp. Anal. Behav.* **31**, 41–51 (1979).
105. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
A seminal work that contributed the key theoretical foundation to research on perceptual decision making.
106. Van Essen, D. C. in *The Visual Neurosciences* (eds Chalupa, L. & Werner, J. S.) 507–521 (MIT Press, Cambridge, Massachusetts, 2004).
107. Van Essen, D. C. *et al.* An integrated software suite for surface-based analyses of cerebral cortex. *J. Am. Med. Inform. Assoc.* **8**, 443–459 (2001).

Acknowledgements

We thank several colleagues who provided critical comments and suggestions during the preparation of this review: S. Baccus, C. Fiorillo, B. Linkenhoker, J. Reppas, A. Rorie and M. Shadlen. We also thank R. Gallistel of Rutgers University who taught us a great deal about matching behaviour and helped us design the oculomotor matching task used in our own studies. L.P.S. was supported by a Stanford Graduate Fellowship and by the Medical Scientist Training Program at the Johns Hopkins University School of Medicine. G.S.C. was also supported by a Stanford Graduate Fellowship and is currently supported by a National Research Service Award predoctoral fellowship from the National Institute of Mental Health. W.T.N. is an Investigator of the Howard Hughes Medical Institute and is also support by the National Eye Institute.

Competing interests statement

The authors declare no competing financial interests.

Online links

FURTHER INFORMATION

SumsDB: http://sumsdb.wustl.edu:8081/sums/
CARET and Surface-Based Atlases: http://brainvis.wustl.edu/caret/

Access to this interactive links box is free online.

© 2005 Nature Publishing Group