The Neural Substrates of Reward Processing in Humans: The Modern Role of fMRI

SAMUEL M. MCCLURE, MICHELE K. YORK, and P. READ MONTAGUE

Experimental work in animals has identified numerous neural structures involved in reward processing and reward-dependent learning. Until recently, this work provided the primary basis for speculations about the neural substrates of human reward processing. The widespread use of neuroimaging technology has changed this situation dramatically over the past decade through the use of PET and fMRI. Here, the authors focus on the role played by fMRI studies, where recent work has replicated the animal results in human subjects and has extended the view of putative reward-processing neural structures. In particular, fMRI work has identified a set of reward-related brain structures including the orbitofrontal cortex, amygdala, ventral striatum, and medial prefrontal cortex. Moreover, the human experiments have probed the dependence of human reward responses on learned expectations, context, timing, and the reward dimension. Current experiments aim to assess the function of human reward-processing structures to determine how they allow us to predict, assess, and act in response to rewards. The authors review current accomplishments in the study of human reward processing and focus their discussion on explanations directed particularly at the role played by the ventral striatum. They discuss how these findings may contribute to a better understanding of deficits associated with Parkinson's disease. NEUROSCIENTIST 10(3):260–268, 2004. DOI: 10.1177/1073858404263526

KEY WORDS *Reward, fMRI, Atriatum, Amygdala, Orbitofrontal cortex, Parkinson's disease*

The past several years have witnessed a tremendous advance in the study of human brain reward processing using fMRI. The strength and consistency of the findings have established fMRI blood oxygenation level–dependent (BOLD) signal imaging as a viable and powerful method for studying reward-related brain processes in people. In this article, we address these recent findings to highlight what is currently known regarding how the human brain senses rewards, learns to predict rewards, and generates actions to acquire rewards.

The capacity to seek rewards as goals is essential for the survival and reproduction of all mobile creatures. Rewards are defined operationally as those stimuli that positively reinforce behavior; that is, rewards increase the probability of a behavior. Stated differently, when one experiences a reward, one becomes more likely to engage in behaviors that lead to the reward in the future. Food, water, and sexual stimuli are called primary rewards because they reinforce behaviors without having to be learned. Other stimuli, such as money, gain reward value by learned association with primary rewards. Indeed, almost any sensory stimulus can acquire reward value through appropriate conditional learning.

One active area of research is the question of how diverse types of rewards map onto brain function. The classic definition used here is a relatively crude notion, and there is no reason to think that the brain responds equivalently to all types of reward. Regardless, reward processing can be broken down into necessary computational steps that can then be addressed experimentally.

Identification of Brain Reward Circuitry

The first step in establishing fMRI as a valid technique in the study of reward processing was to determine whether rewards induce a measurable BOLD signal and, if so, which brain structures are recruited when subjects experience rewards of different modalities. To this end, a wide range of rewarding stimuli have now been shown to modulate BOLD signal responses. The brain areas activated vary with respect to the behavioral task; however, rewarding stimuli will consistently increase activity in a common set of neural structures that include the orbitofrontal cortex (OFC), amygdala, and ventral striatum/nucleus accumbens (NAc) (see below). Areas of the prefrontal cortex and anterior cingulate gyrus have also been observed in a number of studies (Knutson and

From the Human Neuroimaging Lab, Center for Theoretical Neuroscience, Division of Neuroscience (SMM, PRM), and the Department of Neurosurgery (MKY), Baylor College of Medicine, Houston, Texas. Samuel M. McClure is now at the Department of Psychology, Princeton University, Princeton, New Jersey.

This work was supported by grants from the Kane Family Foundation (PRM) and the National Institutes for Health RO1 MH52797 (PRM), RO1 DA11723 (PRM).

Address correspondence to: Samuel M. McClure, Department of Psychology, Green Hall, Princeton University, Princeton, NJ 08544 (e-mail: smcclure@princeton.edu).

Copyright © 2004 Sage Publications
ISSN 1073-8584

others 2003; Sanfey and others 2003; Ullsperger and Von Cramon 2003). These regions seem to be more involved, and still little understood, in the complex process of integrating reward information for the purpose of action selection. We restrict our focus in this article to brain areas related to sensing, predicting, and valuing rewards.

Activity changes in the OFC-amygdala-NAc reward circuit are contingent on the receipt of a wide range of rewarding stimuli. The rewards employed in fMRI experiments have included primary rewards such as fruit juice and water (Berns and others 2001; O'Doherty, Rolls, and others 2001; O'Doherty and others 2002; Pagnoni and others 2002; McClure, Berns, and others 2003; O'Doherty, Dayan, and others 2003), appetitive smells (Gottfried and others 2002; Anderson and others 2003), sexual stimuli (erotic movies; Arnow and others 2002), and sexual behavior (Komisaruk and others 2002). Conditioned rewards have also been used. A conditioned reward is not initially rewarding but accrues its reward status through some type of conditioning, that is, through the consistent temporal pairing of the stimulus with some form of primary reward. Commonly used conditioned rewards, including money and positive feedback (Delgado and others 2000; Elliott, Friston, and others 2000; Knutson and others 2000, 2001; O'Doherty, Kringelbach, and others 2001), also activate these same neural structures. Arbitrary stimuli such as abstract visual cues (Knutson and others 2000, 2001; Gottfried and others 2002; O'Doherty and others 2002) and brief flashes of light (Pagnoni and others 2002; McClure, Berns, and others 2003) that have been conditioned to predict the occurrence of rewards reveal similar patterns of activation. Even social rewards, such as beautiful faces (Aharon and others 2001), social interactions (Rilling and others 2002), affect-laden words (Hamann and Mao 2002), and pleasant touch (Rolls and others 2003), recruit this same coterie of neural structures.

The common pattern of activation found to respond to these diverse stimuli has led to the hypothesis that the brain may process rewards along a single final common pathway in the form of a kind of common neural currency. Functionally, this may reveal an important insight into how reward information is processed in the brain: a common network allows widely different rewards to be directly compared for the purpose of choosing between possible courses of action (Shizgal 1997; Montague and Berns 2002).

Although fMRI has been successful in identifying the neural structures that respond to rewards, less progress has been made in determining what roles these different structures play in reward processing. For the remainder of this article, we highlight some of the progress that has been made for different brain structures. We then focus specifically on the ventral striatum, for which the most detailed hypothesis exists. Finally, we speculate on what importance these findings may have in terms of understanding a subset of the motor and cognitive deficits resulting from Parkinson's disease.

Function of Reward-Related Neural Structures as Revealed with fMRI

fMRI is still a new technique for analyzing brain function; the initial proof-of-concept experiments were conducted only a little more than a decade ago (Ogawa and others 1990; Belliveau and others 1991). It was some time following the development of fMRI before statistical and methodological techniques were developed that allowed for the investigation of changes in neural activity on functionally relevant time scales (Rosen and others 1998). The resultant methodology, event-related fMRI, has been essential for parsing separate functions of different reward-related brain structures. This direction of investigation is still in its infancy. Regardless, progress in the understanding of some areas has been substantial, especially where it has built on findings from animal experiments.

OFC Reward Valuation

In terms of connectivity, the OFC is in a unique location with regard to reward processing. It receives direct inputs from primary taste and olfactory cortices as well as from higher-order visual and somatosensory areas (Elliott, Dolan, and others 2000; Rolls 2000). The OFC is therefore in an ideal location for storing the reward value of sensory stimuli. In fact, experiments in rats have revealed that OFC neurons respond preferentially to different tastes (Rolls 2000). Furthermore, the degree to which taste stimuli activate OFC neurons is related to the stimulus reward value (i.e., it is significantly reduced in the case of satiety) (Rolls 2000). Experiments in macaque have revealed similar findings, with the additional discovery that the amplitude of induced neural activity indicates the relative reward value of stimuli compared to other available rewards (Tremblay and Schultz 1999).

In humans, the function of the OFC is best known from the infamous case of Phineas Gage, who sustained substantial damage to the OFC and parts of the prefrontal cortex. As a result of his injuries, Gage uncharacteristically began to behave inappropriately in social contexts and was prone to bad business judgments (Damasio 1994). Modern studies of patients with more specific damage to the OFC have indicated that although patients can accurately assess the relative value of actions, they fail to behave appropriately according to this information (Bechara and others 1994; Rolls and others 1994). For example, OFC patients fail to suppress incorrect responses in simple decision-making tasks after reward contingencies have been switched, even though they know the appropriate response (Rolls and others 1994).

fMRI has revealed that the OFC responds in conditions that require both approach behavior and response inhibition. The location of the OFC activity differs such that rewarded actions (approach) tend to activate more medial areas, and punished actions (response inhibition) tend to activate more lateral areas (Elliott, Dolan, and

others 2000). Recent studies challenge this simple medial-lateral distinction (O'Doherty, Critchley, and others 2003). However, although it is still unclear how different areas of the OFC correlate with the valence of reinforcing stimuli and actions, there is little doubt that the OFC participates in an important mapping between sensory stimuli and reward value (Montague and Berns 2002). Furthermore, this mapping seems vital for generating appropriate reward-directed behavioral responses.

Amygdala Reinforcer Intensity

The amygdala is well recognized for its association with negative emotions and fear. Lesions to the amygdala reduce aggressive behavior and fear responses in monkeys (Weiskrantz 1956). Amygdala function is also necessary for fear conditioning (LeDoux 2000). fMRI studies have confirmed that amygdala activity is related to sensing aversive stimuli. For example, the amygdala is preferentially activated by faces exhibiting fearful or angry expressions (for review, see Calder and others 2001). However, the notion that the amygdala is involved only in responding to aversive stimuli is now under question (Baxter and Murray 2002), in part due to evidence generated using fMRI. In particular, it has been consistently found that the amygdala is recruited following the presentation of pleasurable, positively reinforcing stimuli (Hamann and Mao 2002; Anderson and others 2003; Hommer and others 2003; Small and others 2003). Furthermore, when responses are directly compared for rewarding and aversive stimuli, the amygdala seems to relate to how arousing the stimuli are as opposed to the valence of the stimuli (positive versus negative) (Fig. 1; Anderson and others 2003; Small and others 2003). The older findings that seemed to definitively link the amygdala to negative emotions have been reinterpreted to result from the fact that aversive stimuli are generally more salient than rewarding stimuli (Anderson and others 2003). Thus, the fact that the amygdala seems more correlated with aversive events may be indicative of the stronger behavioral relevance of negative versus positive stimuli.

Ventral Striatum (Nucleus Accumbens) Reward Prediction

The ventral striatum has generally received the most attention in the study of reward processing due to its long history of association with rewards and the mesencephalic dopamine system. Animal studies have revealed that electrical stimulation of the ventral striatum, or administration of dopamine receptor agonists to the ventral striatum, is highly rewarding (Ikemoto and Panksepp 1999). Evidence that systemic administration of a pharmacological agent enhances dopamine release in the ventral striatum is frequently taken as evidence of the agent's potential addictiveness (Koob and Bloom 1988). Likewise, the ventral striatum has been commonly observed in fMRI studies of reward processing with BOLD signal changes found to scale with reward amplitude (Knutson and others 2001; Gottfried and others 2003). The timing of BOLD signal increases alternates between the delivery of reward (Delgado and others 2000; Elliott, Friston, and others 2000) and stimuli that predict the occurrence of reward (i.e., signal reward anticipation; Knutson and others 2001; O'Doherty and others 2002). All these findings are accounted for with the hypothesis that the ventral striatum responses signal errors in the prediction of reward (Montague and others 1996; Schultz and others 1997; Berns and others 2001; Pagnoni and others 2002; McClure, Berns, and others 2003; O'Doherty, Dayan, and others 2003). We discuss this theory in detail below and detail experiments that test the theory directly.

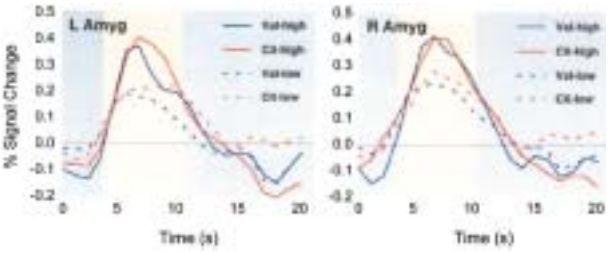

Fig. 1. Amygdala activity is related to stimulus intensity. Subjects were presented with different concentrations of pleasant (citral [Cit]) and unpleasant (valeric acid [Val]) odors. Responses in the left and right amygdala correlated with the concentration of the odorant but were independent of valence (pleasant or unpleasant). This, and other studies, has led to a reevaluation of previous findings that implicated the amygdala as strictly related to negative or aversive stimuli. Adapted from Anderson and others (2003).

In all, several important stages of reward processing have been identified and roughly localized in the brain. The occurrence of a salient reinforcing stimulus may be signaled by the amygdala, which then may trigger a learning signal in the ventral striatum so that the reward may become better predicted in the future (formation of stimulus-reward association). The value of the reward may then be assessed in the OFC, to be used by the prefrontal cortex to decide a course of action consistent with current goals. This is a very tentative description and is almost certainly incorrect in important ways. Nevertheless, fMRI has provided considerable evidence in support of this general view, which speaks directly to the power of the technique.

Prediction Error Theory of Ventral Striatum Activity

Reward processing posits a very specific computational challenge to the nervous system: generate actions to efficiently acquire rewards as necessary for survival. Although numerous brain structures have now been identified with different stages of reward processing, precisely how they interact to generate reward-directed behaviors is unclear. That is, we are far from being able to produce an autonomous robot based on principles derived from our understanding of brain reward structures. Largely, this is symptomatic of a lack of specifici-

ty in hypotheses for what the different brain structures are computing. One exception to this is the dorsal and ventral striatum, where BOLD signal changes have been hypothesized to reflect a prediction error signal that is important for learning to predict rewards and to bias actions that seek rewards (McClure, Berns, and others 2003; McClure, Daw, and others 2003).

The prediction error hypothesis is based on reinforcement learning models from computer science (Sutton and Barto 1998). In short, reinforcement learning seeks to generate algorithms for learning appropriate actions based solely on the intermittent receipt of rewards and punishments. This approach may be contrasted with computer programs that rely on expert programmers to code decision-making algorithms that are then followed without exception. In a complicated environment, there is no way to anticipate all possible environmental scenarios, and so hard-coded solutions are doomed to failure at some point.

Solutions to the reinforcement learning problem have begun with the assumption that there exists a value function, V^*, that relates the expected total reward that is available from any state, s, in the world into the distant future:

$$V^*(s_t) = E\{r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots\},$$

where r_t, r_{t+1}, . . . index the occurrence of rewards at the current time, t, and at subsequent time steps into the future ($t + 1$, $t + 2$, . . .). This value function takes into account the stochastic nature of the world by taking the expected value of these future rewards, E. In addition, more immediate rewards are given higher value than later rewards through the scale factor $\gamma (0 \leq \gamma \leq 1)$.

If an organism could learn $V^*(s_t)$ perfectly for all s_t, then knowing how to navigate through the world would be solved: Take actions that take you to states with greatest value $V^*(s_{t+1})$. Reinforcement learning is reduced, then, to learning a good approximation to V^*. One method for learning such an approximation is known as temporal difference learning and takes advantage of an inherently recursive relationship in V^* (Bellman 1957):

$$V^*(s_{t+1}) = E\{r_{t+1} + \gamma r_{t+1} + \gamma^2 r_{t+3} + \ldots\}$$
$$\Rightarrow V^*(s_t) = E\{r_t + \gamma V^*(s_{t+1})\}.$$

The quality of an estimate (V) of the true valuation function is given by its ability to account for all future rewards and hence by how well it satisfies the Bellman equation through time. The degree to which this fails is captured by the error equation

$$\delta(t) = r(t) + \gamma V(t+1) - V(t).$$

This error, $\delta(t)$, is referred to as the *prediction error* and can be used as a learning signal. When δ is positive, it indicates that the estimation $V(t)$ was too pessimistic and hence that future predictions should be greater. The inverse relation holds for negative values of δ. This learning procedure can be shown to lead to optimal behavior for a wide range of environments (Sutton and Barto 1998).

Dopamine neurons have been proposed to generate a prediction error signal akin to $\delta(t)$ (Montague and others 1996). Single neuron recordings from dopamine neurons in monkeys during simple behavioral tasks have strongly supported this hypothesis (Montague and others 1996; Schultz and others 1997). If these changes in rate of action potential generation translate into changes in BOLD response, then similar changes may be hypothesized to occur at the location of dopaminergic cell bodies (in the ventral midbrain) and in areas of dense dopaminergic innervation (including the dorsal and ventral striatum and ventral prefrontal cortex). Several studies have now been performed to test this hypothesis directly.

In one experiment, primary rewards (juice and water) were delivered to subjects separately in predictable and unpredictable sequences (Berns and others 2001). According to the prediction error hypothesis, the evoked neural response under these two conditions should be vastly different. Rewards delivered during predictable sequences (juice and water delivery was alternated and given at a fixed interstimulus time interval) should decrease to zero as subjects learn to predict the onset of stimulus delivery. That is, juice and water delivery occur exactly at the expected time and in the expected amount, so there is no error in predicted reward at the time of their receipt. The expected change in activity is very different for rewards delivered in unpredictable sequences (random order of juice and water at a random interstimulus time interval). In an unpredictable sequence, subjects can learn neither what reward is coming next nor when it is coming. Therefore, each instance of reward delivery should represent more than the subjects could expect at the delivery time and so should evoke a positive prediction error signal ($\delta(t) > 0$). The BOLD signal data bore this hypothesis out. The change in activity following rewards delivered in the unpredictable sequence were significantly greater than the change following the same stimuli delivered in the predictable sequence in the ventral striatum and ventral prefrontal cortex (Fig. 2; Berns and others 2001).

In a second set of experiments, a remarkable consequence of the prediction error hypothesis was tested for how BOLD signals are predicted to change in the course of a conditioning experiment (McClure, Berns, and others 2003; O'Doherty, Dayan, and others 2003). Consider the experiment diagrammed in Figure 3. In this experiment, a subject is trained to expect the delivery of a reward at a fixed time following a conditioned stimulus (CS). The CS is selected to be motivationally neutral, so that when the subject observes it, no reward is perceived and there is accordingly no change in dopaminergic cell firing before training (Fig. 3*B*, panel 1). Also, before training, the reward cannot be expected, so its delivery is initially unpredictable and results in a positive prediction error response (Figs. 3*A* and 3*B*, panel 1). As training

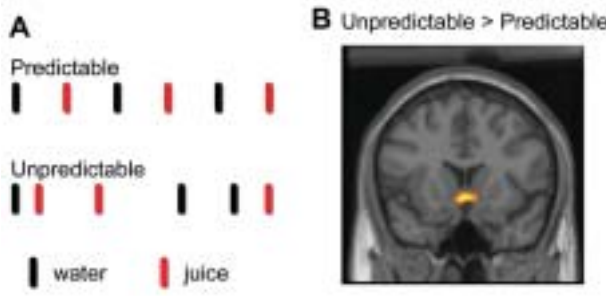

Fig. 2. Activity in the ventral striatum is dependent on the predictability of reward delivery. The neural response to fruit juice (red bars) and water (black bars) can be modulated by delivering the stimuli in either predictable (regular order and timing) or unpredictable (irregular order and timing) sequences. Juice and water delivered during unpredictable sequences elicits much greater activity in the ventral striatum (including nucleus accumbens) compared with predictable sequences. This finding is consistent with the hypothesis that activity in the ventral striatum reflects a learning signal important for acquiring stimulus-reward contingencies. Adapted from Berns and others (2001).

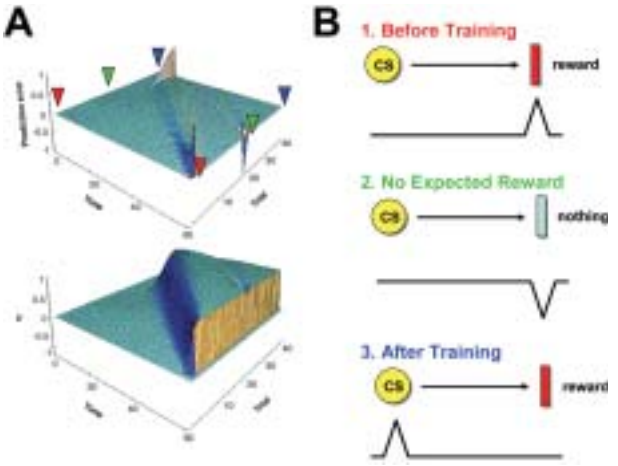

Fig. 3. Reward prediction errors during a simple conditioning experiment predicted by reinforcement learning. The expected prediction error signal through time is plotted for each of several repeated trials (*A*, *top*) along with the learned value of the world (*V; B*, *bottom*). Before the model has learned to expect the reward following a conditioned stimulus (CS), the reward delivery is unexpected and elicits a positive prediction error (*A*, red triangles; *B*, before training). As the CS → reward pairing is repeated, the CS comes to predict the occurrence of reward. This renders the reward delivery predictable (no associated prediction error) and causes the prediction error signal to shift to the time of the CS (*A*, blue triangles; *B*, after training). When the reward is unexpectedly withheld, the absence of reward at the predicted time evokes a negative prediction error signal (*A*, green triangles; *B*, no expected reward). Each of these responses has been observed in recording from dopamine neurons in monkeys (Schultz and others 1997) and in BOLD signal measures from human ventral striatum (Fig. 4; McClure, Berns, and others 2003; O'Doherty, Dayan, and others 2003). Adapted from Montague and others (1996).

continues, the responses to the reward and CS are predicted to change dramatically. The subject should come to predict the onset of the reward. Thus, the amount of reward received should be precisely what is expected, and the reward delivery should elicit no prediction error. Instead, the prediction error is hypothesized to shift to the time of the CS, as the onset time of the CS is random (unpredictable) and carries with it a prediction of increased future reward. The degree of learning can be further assessable after training by omitting the delivery of the reward. In this case, the absence of juice delivery represents a negative prediction error ($\delta(t) < 0$), indicating that less reward was received than was expected. These predictions are fully met in the response of dopamine neurons (Schultz and others 1997). fMRI experiments indicate that BOLD signal changes in the ventral striatum also possess these characteristics (Figs. 4 and 5; McClure, Berns, and others 2003; O'Doherty, Dayan, and others 2003).

In addition to these two experiments, and as discussed above, the prediction error theory provides a parsimonious account of nearly all the imaging results that pertain to the ventral striatum (but see Becerra and others 2001). Given that such a prediction error signal is well suited to function as a learning signal, the hypothesis naturally follows that the ventral striatum may be a locus for learning stimulus-reward associations. However, reinforcement learning methods offer an additional use for such prediction error signals. Temporal difference learning networks are well suited for generating signals to be used for biasing action selection (McClure, Daw, and others 2003). Therefore, chronically reduced prediction error signaling may manifest itself in two ways: 1) as a reduced ability to adaptively learn stimulus-reward associations and 2) as a reduced ability to bias action selections. In the final section of this article, we briefly discuss how this interpretation provides a simple account for both the motor and cognitive deficits seen in Parkinson's disease.

Reinterpretation of Parkinsonian Motor and Cognitive Deficits

Parkinson's disease is a neurodegenerative motor disorder characterized by a clinical triad of akinesia, rigidity, and tremor. In addition, many Parkinson's disease patients also demonstrate significant postural instability and levodopa-induced dyskinesias. Parkinson's disease patients demonstrate loss of neurons in the pars compacta of the substantia nigra, with motor deficits apparent after 80% to 90% cell loss. The resultant disturbed function of the dopamine neurons in the nigro-striatal pathway underlies the motor deficits (Alexander and others 1986; Starr and others 1998). Parkinson's disease–associated cognitive deficits are also evident even at early stages of disease progression (Starkstein and others 1989; Brown and Marsden 1990; Cooper and others 1991; Dubois and Pillon 1997; Piccirilli and others 1997; Woods and others 2003). Dopamine depletion in the lateral orbitofrontal and the dorsolateral prefrontal circuits has been suggested as a possible mechanism of

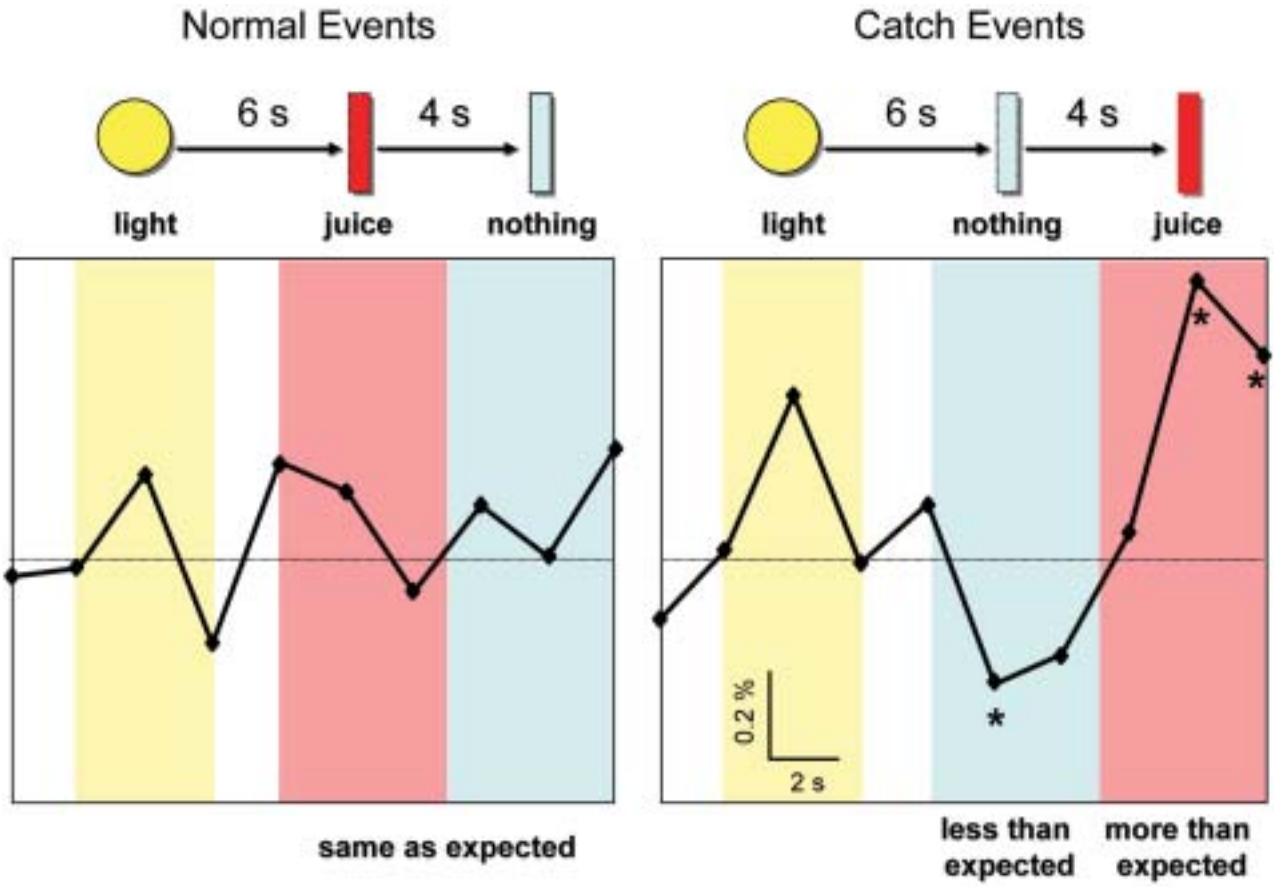

Fig. 4. fMRI blood oxygenation level–dependent (BOLD) responses in the left striatum during a simple conditioning experiment. Subjects were trained to expect juice delivery at a fixed time (6 sec) following a flash of light. The response to these pairings (normal events) after training is shown on the left. Allowing for expected lag (~2 sec) and duration (4–6 sec) of the hemodynamic response, the response to the predictive light cue is slightly elevated from baseline (time points indicated in yellow), but there is no response to the juice delivery (time points indicated in red) or during a control period following juice delivery (indicated in blue). When the delivery of juice is delayed after training, two prediction errors are expected to occur. At the time of expected juice, less reward is received than expected and hence a negative prediction error occurs. This event correlates with a decreased BOLD response ($P < 0.05$) in the left striatum. Juice delivery at the delayed time is more reward than could be predicted by training and hence should elicit a positive prediction error. Accordingly, increased BOLD signal is found. Adapted from McClure, Berns, and others 2003.

cognitive impairment in Parkinson's disease (Alexander and others 1986; Starr and others 1998). Dysfunction in these neural circuits results in cognitive deficits similar to patients with a frontal lobe syndrome (Taylor and others 1990; Woods and others 2003).

The computational role ascribed to dopamine neurons described above offers an account of these motor and cognitive symptoms. To begin with the motor deficits, reinforcement learning theory suggests that reward prediction errors may be used to bias action selection (Fig. 6; Montague and others 1996; McClure, Daw, and others 2003). Animal research has shown that dopamine release precedes (Phillips and others 2003) and is necessary (Berridge and Robinson 1998) for the initiation of reward-directed behavior. In the case of Parkinson's disease, the damaged dopamine system would result in a decreased ability to bias action selection. The resulting affect may appear both as indecision in action selection (tremor) and as the inability to initiate desired behavior (rigidity, akinesia).

Reinforcement learning theory also provides an explanation for the frontostriatal cognitive changes associated with Parkinson's disease, including set shifting, initiation, inhibition, problem solving, and decision making. Similar to the motor system, the reduction in dopamine in the basal ganglia-thalamocortical loop might also modify action selection during cognitive tasks. The resulting cognitive deficits may manifest as the inability to initiate a desired cognitive change (e.g., set shifting, initiation, inhibition). Temporal difference learning may also account for problem-solving and decision-making deficits in Parkinson's disease through dopamine's direct and indirect role in learning value predictions. A dysfunction in the dopaminergic system would suppress the desirability of one action over another and thus result in a reduction in the probability of making or changing any

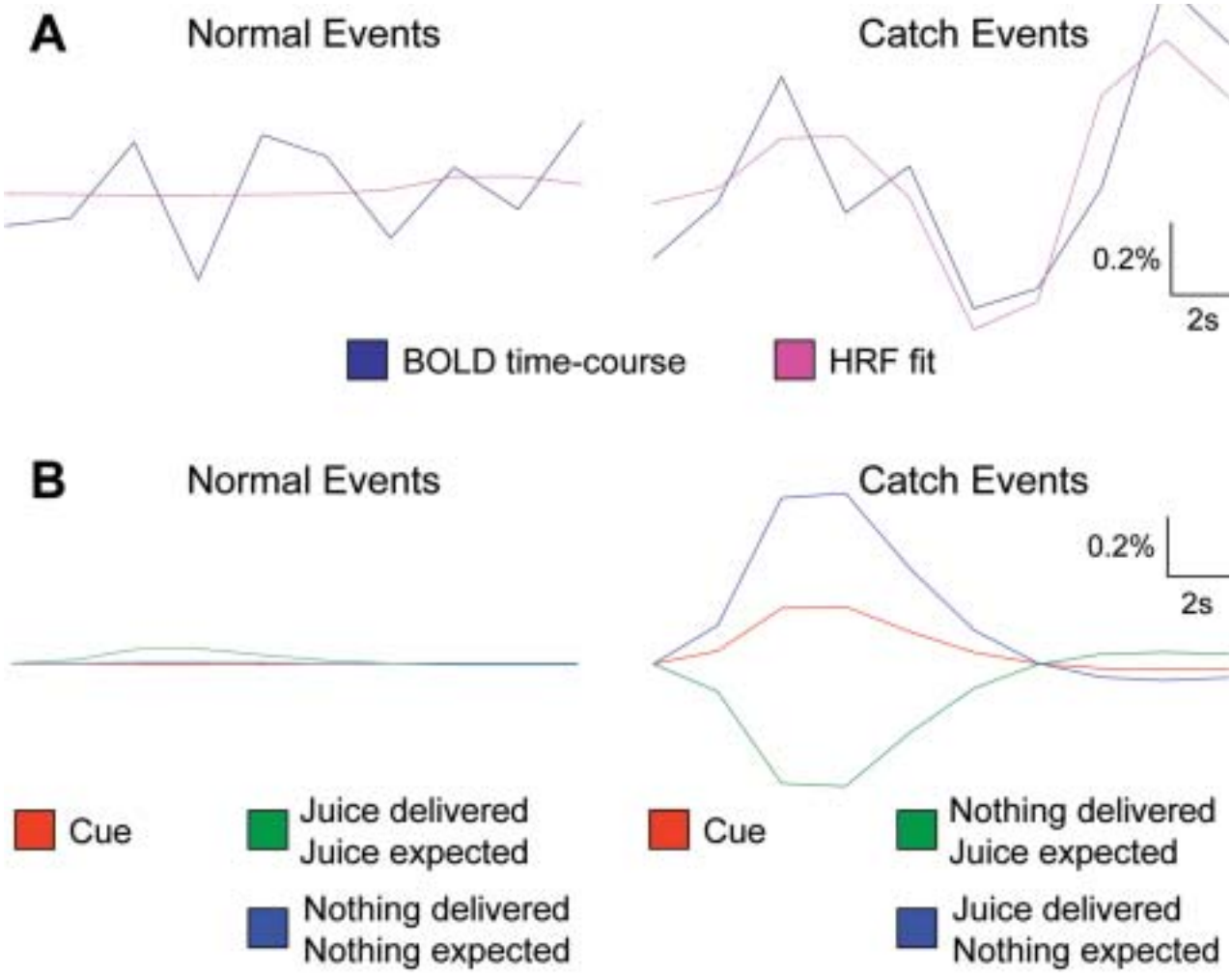

Fig. 5. Analysis of striatal response during conditioning experiment using generic hemodynamic response function. *A*, The mean blood oxygenation level–dependent (BOLD) time-courses for normal and catch events (blue curves; see Fig. 4) were fit with a series of hemodynamic response functions (HRFs) convolved with the event times. Fitting the HRFs to produce the least squared fit to the data produces the curves shown in magenta. For catch events, this procedure accounted for 72% of the variance in the data ($r^2 = 0.72$). *B*, Plotting the individual HRFs that produce the best fits in (*A*) reveals that juice delivered at unexpected times in catch events (catch events, blue trace) elicits a greater change in BOLD response than juice delivered at expected times in normal events (normal events, green trace). Furthermore, the absence of juice delivery at expected times (catch events, green trace) causes a significant decrease in BOLD signal relative to the absence of juice delivery at control times (normal events, blue trace).

decision (e.g., perseveration) and/or increasing the time required to solve a problem (e.g., bradyphrenia).

The severity and progression of motor and cognitive deficits in Parkinson's disease vary across patients. A subset of Parkinson's disease patients experience tremor-dominant Parkinson's disease without significant rigidity and akinesia, whereas others experience significant and debilitating rigidity and akinesia with only mild tremor. Similarly, a subset of Parkinson's disease patients do not demonstrate cognitive deficits consistent with frontostriatal dysfunction, whereas others begin to experience subtle deficits in set shifting, initiation, and inhibition early in the disease progression. Parkinson's disease patients who demonstrate significant rigidity and akinesia with frontostriatal cognitive decline may constitute a distinct clinical category of Parkinson's disease patients. These patients may provide additional support for the reinterpretation of Parkinson's disease–related motor and cognitive changes using a computational prediction error theory approach.

The best current treatment for Parkinson's disease employs levadopa to increase basal levels of dopamine in the brain. What is clear from animal and clinical research, and now also from fMRI, is that the timing of dopamine release is critical for proper dopaminergic function. Transient increases in dopamine are necessary for both selectively motivating actions and performing important cognitive tasks. Improved understanding of the computational role of dopaminergic transmission in reward processing, aided by fMRI, may provide critical insights into Parkinson's disease that may facilitate improved treatments.

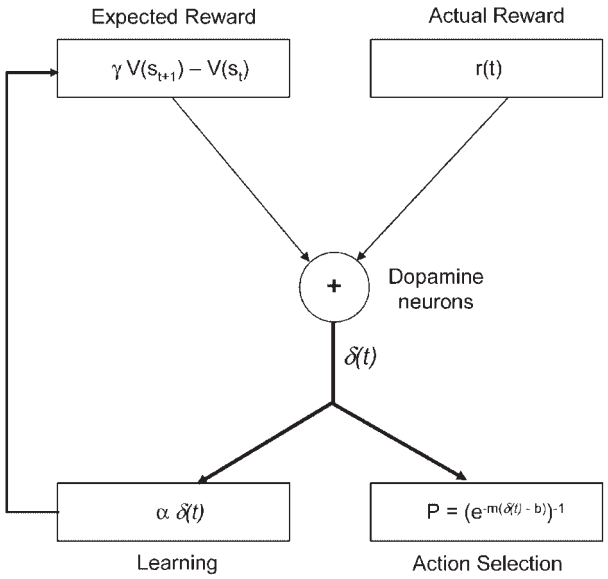

Fig. 6. Reward prediction errors are used both in reward learning and in reward-based decision making. The reinforcement learning model generates reward prediction error signals by comparing expected and actual rewards. This difference, the reward prediction error ($\delta(t)$), indicates the magnitude and direction of learning that must occur to improve performance in the world. If this same circuit is used offline and driven by internally generated representations of possible behavioral acts, then the prediction error at any moment indicates the relative value of the currently considered act. The prediction error in this case may be used to bias action selection. Adapted from McClure, Daw, and others (2003).

Looking Forward

Animal studies of reward processing have been an area of intense investigation for more than 50 years. With the advent of fMRI and event-related methods, the study of reward processing in humans has now become a reality. This opens the door to studies of how reward-related information is used in decision making and other complex cognitive tasks that in animals were difficult or impossible due to challenges posed by training. In only a few years, fMRI has proven itself to be a powerful technique in the study of reward processing and has contributed to our understanding of the function of numerous reward-related brain structures. These new findings may provide crucial insights into disease-related deficits and may contribute to improved treatments.

References

Aharon I, Etcoff N, Ariely D, Chabris CF, O'Donner E, Breiter HC. 2001. Beautiful faces have variable reward value: fMRI and behavioral evidence. Neuron 32:527–51.

Alexander GE, DeLong MR, Strick PL. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357–81.

Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, and others. 2003. Dissociated neural representations of intensity and valence in human olfaction. Nature Neurosci 6:196–202.

Arnow BA, Desmond JE, Banner LL, Glover GH, Solomon A, Polan ML, and others. 2002. Brain activation and sexual arousal in healthy, heterosexual males. Brain 125:1014–23.

Baxter MG, Murray EA. 2002. The amygdala and reward. Nat Rev Neurosci 3:563–73.

Becerra L, Breiter HC, Wise R, Gonzalez RG, Borsook D. 2001. Reward circuitry activation by noxious thermal stimuli. Neuron 32:927–46.

Bechara A, Damasio AR, Damasio H, Anderson SW. 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50:7–15.

Belliveau JW, Kennedy DN, McKinstry RC, Buchbinder BR, Weisskoff RM, Cohen MS, and others. 1991. Functional mapping of the human visual cortex by magnetic resonance imaging. Science 254:716–9.

Bellman RE. 1957. Dynamic programming. Princeton (NJ): Princeton University Press.

Berns GS, McClure SM, Pagnoni G, Montague PR. 2001. Predictability modulates human brain response to reward. J Neurosci 21:2793–8.

Berridge KC, Robinson TE. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Rev 28:309–69.

Brown RG, Marsden CD. 1990. Cognitive function in Parkinson's disease: from description to theory. Trends Neurosci 13:21–9.

Calder AJ, Lawrence AD, Young AW. 2001. Neuropsychology of fear and loathing. Nat Rev Neurosci 2:352–63.

Cooper JA, Sagar HJ, Jordan N, Harvey NS, Sullivan EV. 1991. Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. Brain 114:2095–122.

Damasio AR. 1994. Descartes' error: emotion, rationality and the human brain. New York: Putnam.

Delgado MR, Nystron LE, Fissell C, Noll DC, Fiez JA. 2000. Tracking the hemodynamic responses to reward and punishment in the striatum. J Neurophysiol 84:3072–7.

Dubois B, Pillon B. 1997. Cognitive deficits in Parkinson's disease. J Neurol 244:2–8.

Elliott R, Dolan RJ, Frith CD. 2000. Dissociable functions of the medial and lateral orbitofrontal cortex: evidence from neuroimaging studies. Cereb Cortex 10:308–17.

Elliott R, Friston KJ, Dolan RJ. 2000. Dissociable neural responses in human reward systems. J Neurosci 20:6159–65.

Gottfried JA, O'Doherty J, Dolan RJ. 2002. Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. J Neurosci 22:10829–37.

Gottfried JA, O'Doherty J, Dolan RJ. 2003. Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301:1104–7.

Hamann S, Mao H. 2002. Positive and negative emotional verbal stimuli elicit activity in the left amygdala. Neuroreport 13:15–9.

Hommer DW, Knutson B, Fong GW, Bennett S, Adams CM, Varner JL. 2003. Amygdalar recruitment during anticipation of monetary rewards. Ann N Y Acad Sci 985:476–8.

Ikemoto S, Panksepp J. 1999. The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. Brain Res Rev 31:6–41.

Knutson B, Adams CM, Fong GW, Hommer D. 2001. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. J Neurosci 21:RC159.

Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D. 2003. A region of the mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. Neuroimage 18:263–72.

Knutson B, Westdorp A, Kaiser E, Hommer D. 2000. FMRI visualization of brain activity during a monetary incentive delay task. Neuroimage 12:20–7.

Komisaruk BR, Whipple B, Crawford A, Grimes S, Kalnin AJ, Mosier K, and others. 2002. Brain activity (fMRI and PET) during orgasm in women, in response to vaginocervical self-stimulation. Society for Neuroscience Abstracts.

Koob GF, Bloom FE. 1988. Cellular and molecular mechanisms of drug dependence. Science 242:715–23.

LeDoux JE. 2000. Emotion circuits in the brain. Annu Rev Neurosci 23:155–84.
McClure SM, Berns GS, Montague PR. 2003. Temporal prediction errors in a passive learning task activation human striatum. Neuron 38:339–46.
McClure SM, Daw ND, Montague PR. 2003. A computational substrate for incentive salience. Trends Neurosci 26:423–8.
Montague PR, Berns GS. 2002. Neural economics and the biological substrates of valuation. Neuron 36:265–84.
Montague PR, Dayan P, Sejnowski TJ. 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–47.
O'Doherty J, Critchley H, Deichmann R, Dolan RJ. 2003. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci 23:7931–9.
O'Doherty J, Dayan P, Friston KJ, Critchley H, Dolan RJ. 2003. Temporal difference models and reward-related learning in the human brain. Neuron 38:329–37.
O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ. 2002. Neural responses during anticipation of a primary taste reward. Neuron 33:815–26.
O'Doherty JP, Kringelbach ML, Rolls ET, Hornak J, Andrews C. 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci 4:95–102
O'Doherty JP, Rolls ET, Francis S, Bowtell R, McGlone F. 2001. Representations of pleasant and aversive taste in the human brain. J Neurophysiol 85:1315–21.
Ogawa S, Lee TM, Kay AR, Tank DW. 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci U S A 87:9868–72.
Pagnoni G, Zink CF, Montague PR, Berns GS. 2002. Activity in human ventral striatum locked to errors of reward prediction. Nat Neurosci 5:97–8.
Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM. 2003. Subsecond dopamine release promotes cocaine seeking. Nature 422:614–8.
Piccirilli M, D'Alessandro P, Finali G, Piccinin G. 1997. Early frontal impairment as a predictor of dementia in Parkinson's disease. Neurology 48:546–7.
Rilling J, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C. 2002. A neural basis for social cooperation. Neuron 35:395–405.
Rolls ET. 2000. The orbitofrontal cortex and reward. Cereb Cortex 10:284–94.
Rolls ET, Hornak J, Wade D, McGrath J. 1994. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol Neurosurg Psychiatry 57:1518–24.
Rolls ET, O'Doherty J, Kringelbach ML, Francis S, Bowtell R, McGlone F. 2003. Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. Cereb Cortex 13:308–17.
Rosen BR, Buckner RL, Dale AM. 1998. Event-related fMRI: past, present, and future. Proc Natl Acad Sci U S A 95:773–80.
Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. 2003. Decision-making in the ultimatum game. Science 300:1755–8.
Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. Science 275:1593–9.
Shizgal P. 1997. Neural basis of utility estimation. Curr Opin Neurobiol 7:198–208.
Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parish T. 2003. Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron 39:701–11.
Starkstein SE, Bolduc PL, Preziosi TJ, Robinson RG. 1989. Cognitive impairments in different stages of Parkinson's disease. J Neuropsychiatry Clin Neurosci 1:243–48.
Starr PA, Vitek JL, Bakay RA. 1998.. Ablative surgery and deep brain stimulation for Parkinson's disease. Neurosurgery 43:989–1013.
Sutton RS, Barto AG. 1998. Reinforcement learning. Cambridge (MA): MIT Press.
Taylor AE, Saint-Cyr JA, Lang AE. 1990. Memory and learning in early Parkinson's disease: evidence for a "frontal lobe syndrome." Brain Cogn 13:211–32.
Tremblay L, Schultz W. 1999. Relative reward preference in primate orbitofrontal cortex. Nature 398:704–8.
Ullsperger M, Von Cramon DY. 2003. Error monitoring using external feedback: specific roles of the habenular comples, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. J Neurosci 23:4308–14.
Weiskrantz L. 1956. Behavioral changes associated with ablation of the amygdaloid complex in monkeys. J Comp Physiol Psychol 49:381–91.
Woods SP, Troster AI. 2003. Prodromal frontal/executive dysfunction predicts incident dementia in Parkinson's disease. J Int Neuropsychol Soc 9:17–24.