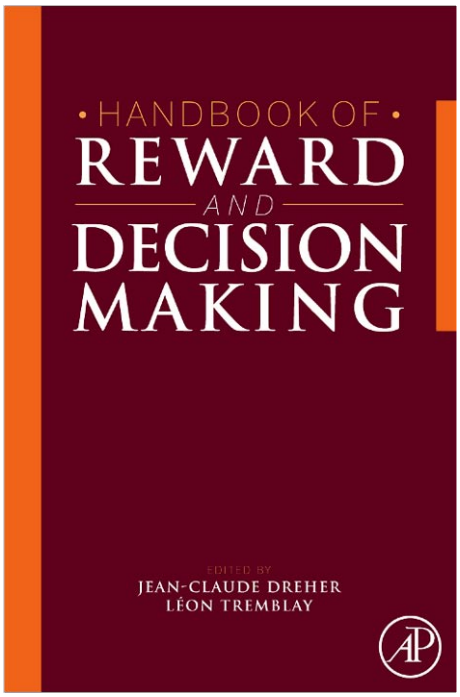




From Jean-Claude her, Decomposing brain signals involved in value-based decision making. In: Dr. Jean-Claude Dreher and Léon Tremblay, editors, Handbook of Reward and Decision Making. Oxford: Academic Press, 2009, pp. 137-163.
ISBN: 978-0-12-374620-7

# Part Two

## fMRI Studies on Reward and Decision Making

# 6 Decomposing brain signals involved in value-based decision making

*Jean-Claude Dreher*

Reward and decision making group, Cognitive Neuroscience Center, CNRS UMR5229, Lyon 1 University 67, Bd Pinel 69675 Bron, France

## Abstract

Deciding between options leading to rewards of different nature depends both upon our internal needs and upon the subjective value of available options. Because there is no single sense organ transducing all types of rewards, our brain may compare different goods using a "common reward currency," allowing us to choose the option with the highest subjective value. Here, we analyze the neural substrates of different value-related signals involved when processing and deciding between rewards of different nature, such as prediction error, uncertainty, subjective value of each option, goal, and decision value. Some of these value signals are computed even without choice, and all of them eventually influence the decision-making process. We review recent neuroimaging work investigating the neural substrates of these different value signals and present recent results showing how genetically-influenced variations in dopamine transmission also influence the response of the reward system in humans.

## Key points

1. Electrophysiological and neuroimaging data show that prediction error is modulated by several factors (probability, magnitude, delay) and is computed regardless of reward type.
2. Reward uncertainty and prediction error are computed in distinct brain networks, and several uncertainty signals co-exist in the brain.
3. Distinct value-based signals (distance between subjective value of different options, goal-value and decision-value) are computed in partially overlapping brain regions.
4. Primary and secondary rewards both activate a common brain network ("common neural currency") and reveal a new functional organization of the orbitofrontal cortex according to reward type (anterior obitofrontal cortex for secondary rewards and posterior orbitofrontal cortex for primary rewards).
5. Genetic variations in dopamine-related genes influence the response of the reward system and may contribute to individual differences in reward-seeking behavior and in predisposition to neuropsychiatric disorders.

## 6.1 Basic computations involved in decision making

When presented with several options, we need to assign subjective values to each of them to make a choice. Based on existing theoretical models of decision making, recent reviews

**Handbook of Reward and Decision Making**

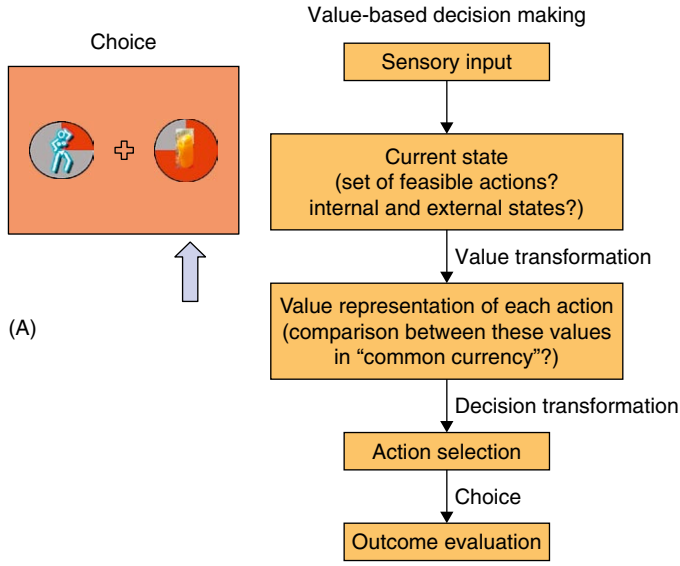


**Figure 6.1** (A) Choosing between options leading to rewards of different nature. Example of a simple choice in which the brain needs to weigh the probability of each option and the nature of potential rewards (here, erotic stimuli or fruit juice) according to internal needs (see paragraph 6.7 for description of fMRI experiment). (B) Basic computations involved in value-based decision making. Based on existing theoretical models of decision making, value-based decision making can be decomposed into basic processes. First, one recognizes the current situation (or state), including internal (e.g., hunger) and external states (e.g., cold), and potential courses of actions (e.g., purchase food). Second, a valuation system needs to weigh available options in terms of cost and benefit (reward/punishment). Third, action selection is implemented based on this valuation. Finally, the chosen action may be re-evaluated based on the actual outcome, leading to updating of the other processes through learning to improve subsequent decisions.

proposed a framework that decomposes value-based decision making into basic processes [1–3] (Fig. 6.1).

The first stage includes a representation of the current situation (or state), including the identification of internal state (e.g., hunger), external state (e.g., cold), and potential courses of actions (e.g., purchase food). Second, a valuation system needs to weigh available options in terms of reward and punishment, as well as cost and benefit. Third, the agent selects an action on the basis of this valuation. Finally, the chosen action may be re-evaluated based on the actual outcome, eventually leading to update the other processes through learning to improve subsequent decisions. Although these processes may occur in parallel, this simplified framework is nonetheless useful to decompose basic computations performed by the brain.

It is still unclear whether there are separate valuation systems in the brain, but a number of studies distinguish between at least two systems: Pavlovian and instrumental conditioning. In Pavlovian (or classical) conditioning, subjects learn to predict outcomes without having the opportunity to act. In instrumental conditioning, animals learn to choose actions to obtain rewards and avoid punishments. Various strategies are possible,

such as optimizing the average rate of acquisition of rewards minus punishments or optimizing the expected sum of future rewards, where outcomes received in the far future are discounted compared with outcomes received more immediately.

The focus of this chapter is to define different value signals that are used by the brain to make a decision and to review our current understanding of their possible neuronal implementation. The brain must perform multiple value computations to make decisions. Even without having to decide between options leading to different rewards, the brain computes a prediction of the value of potential outcomes and compares this prediction with the actual outcome (prediction error signal). This expected value and the reward prediction error signal are modulated by a number of factors, such as the magnitude and probability of reward, the timing uncertainty of the reward delivery, and the delay period between the cue associated with the reward and the outcome delivery. Each of these factors also influence the computation of signals necessary to make a decision, such as goal values, decision values, subjective distance between available options, subjective value of the chosen option, and choice uncertainty.

Below, we review the recent literature on the neural substrates of these different value signals and discuss our own results, showing their contributions to decision making between rewards of different nature or between costly options.

## 6.2 Computing reward prediction error

### *6.2.1 Animal electrophysiology on prediction error*

Prediction errors measure deviations from previous reward expectations. Thus, prediction error can be either positive (when the reward delivered is better than expected) or negative (less or no reward delivered at the expected time) [96, 97]. Prediction errors are used to learn the value of states of the world and are critical for learning how to make better choices in the future. Electrophysiological studies in monkeys indicate that dopaminergic neurons code such a prediction error signal in a transient fashion. This signal may be sent to the striatum and prefrontal cortex to influence reward-dependent learning [4–6].

In classical conditioning experiments, where an association has to be learnt between a visual predictor (conditioned stimulus) and a rewarding outcome (unconditioned stimulus), each of the factors mentioned before (magnitude, probability, timing uncertainty, and delay) influences the phasic prediction signal occurring at the time of the conditioned stimuli. That is, the phasic response of dopamine neurons to the conditioned stimuli monotonically increases with probability and magnitude [7] and decreases with the reward delay in temporal discounting paradigms, both in Pavlovian conditioning [8] (see Chapter 2) and in intertemporal choice [9]. Moreover, at the time of the outcome, the response of dopamine neurons increases with reward delay and magnitude and decreases with increasing reward probability [8,10]. However, the magnitude of activation or suppression of dopaminergic neurons response appears to be identical for different magnitudes that are delivered with maximal uncertainty ($P = 0.5$), despite the fact that the absolute difference between actual and expected volume magnitude varied over a large range [7]. Thus, the responses of dopamine neurons do not appear to scale according to the absolute difference between actual and expected reward. Rather, the sensitivity of the neural responses appears to adapt according to the discrepancy in magnitude between two potential outcomes. Taken together, these results suggest that the dopamine response

reflects the subjective value of the reward and may be sent to a number of neural structures involved in computing value-based signals involved in decision making.

In rodents, recent results also indicate that midbrain dopamine neurons encode decisions for future action [11], as well as the most valuable option in choice situations (reward of shorter delay or larger magnitude), consistently with their proposed role in coding the value of the chosen option post-choice [9]. In this experiment, rats could choose to respond for either a high-value or a low-value reward, and dopamine neurons response reflected the value of the best possible option, independent of which was ultimately selected.

Of particular importance for understanding the functional properties of dopaminergic neurons are two recent findings concerning the roles of the lateral habenula and the globus pallidus (internal segment) in the reward circuitry. The lateral habenula, which provides a key source of input to dopaminergic neurons [12] may suppress the activity of dopamine neurons, inducing pauses in the burst firing of dopamine cells that might be responsible for their negative prediction errors. Lateral habenula neurons respond to conditioned stimuli associated with the absence of reward or the presence of punishment and to punishment itself. They are also inhibited by rewarding outcomes, especially when these are less predictable [13]. Thus, the lateral habenula may control both reward-seeking (associated with the dopaminergic system) and punishment-avoidance behavior (associated with the serotoninergic system), through its projections to these two systems. Moreover, globus pallidus cells may drive the reward-negative responses of lateral habenula neurons [14]. These results help to understand the functional neuroanatomy underlying the response of dopaminergic neurons.

### 6.2.2 Human neuroimaging studies on prediction error

Recent human neuroimaging studies have investigated the neural correlates of the prediction error signal. A number of these studies suggest that activity in the ventral striatum and the prefrontal cortex correlates with prediction errors related to stimulus–response associations or rewards of different types, such as faces, money, or juice [15–21]. When examining the influence of reward magnitude during reward anticipation and at the time of rewarded outcome, increased activity has been observed in several brain regions, particularly in the ventral striatum. For example, increased ventral striatal activation was found with increasing magnitude of anticipated gains but not losses [22,23]. Several studies also investigated the influence of reward probability on brain activation. Some gambling studies found that ventral striatal activity increased with reward probability [19,24,25] while a cued reaction time study failed to find ventral striatal activation as a function of increasing probability [23]. In some of these studies, a region of the medial prefrontal cortex also showed increasing activation during anticipation of rewards with increasing probability [23,25].

In a recent monetary fMRI study using slot machines varying known reward probability and magnitude, we could distinguish between transient and sustained signals using a fixed long anticipatory period [20]. We found that the midbrain was activated both transiently with the prediction error signal and in a sustained fashion with reward uncertainty (see section 6.3). Moreover, distinct activity dynamics were observed in post-synaptic midbrain projection sites: the prefrontal cortex responded to the transient prediction error signal while the ventral striatum covaried with the sustained reward uncertainty signal (Fig. 6.2).

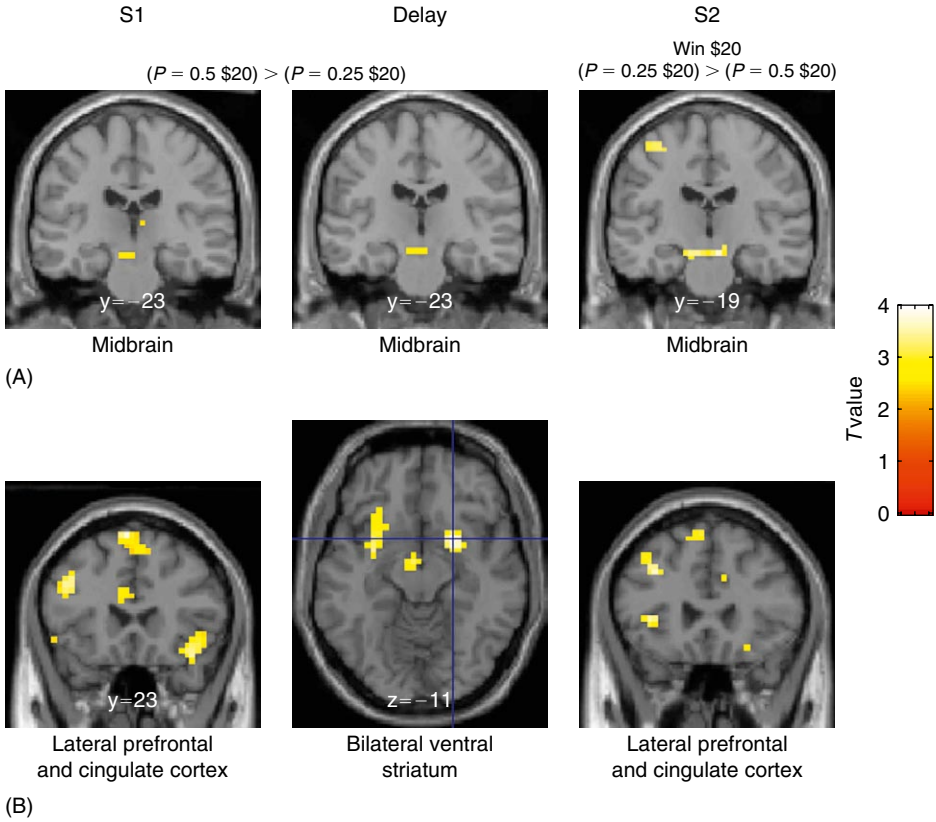


**Figure 6.2** (A) Transient and sustained midbrain activities. Location of transient midbrain responses covarying with the error prediction signal at the cue S1 (*Left*) and at the rewarded outcome S2 (*Right*). Consistent with electrophysiological recordings [10], the human midbrain region was transiently activated with higher reward probability at the cue S1 and with lower reward probability at the rewarded outcome S2. Moreover, the midbrain region showed higher sustained activity with reward uncertainty during the delay [20]. (B) Location of transient midbrain and prefrontal responses covarying with the error prediction signal at the cue S1 (*Left*) and at the rewarded outcome S2 (*Right*). Middle: location of sustained bilateral ventral striatum activites covarying with the reward uncertainty signal during the delay period.

The frontal network we observed both at the time of the cue and at the time of the outcome was specifically involved with the reward prediction error signal because it was not significantly activated by reward uncertainty during the delay, and was significantly more activated in association with these phasically-modeled responses than in association with a sustained-modeled response related to reward uncertainty during the delay period. Our results extend previous fMRI reports that the dorsolateral prefrontal cortex, inferior frontal gyrus, and orbitofrontal cortex activity correlates with a prediction error signal related to abstract stimulus–response associations or taste reward, although some of these studies focused more on ventral striatal activity [15–21]. The lateral prefrontal

cortex may generate the reward prediction because neurons from this brain region represent predictions about expected rewards according to the context [26,27].

In two recent fMRI studies, we then investigated how prediction error is modulated not only by reward probability and magnitude but also by reward type (money, fruit juice, and erotic stimuli) and by reinforcement nature (reward *versus* punishment). In a first study, we explicitly informed subjects on subsequent reward type (erotic stimuli or monetary reward), probability, and intensity. We found that activity in the ventral striatum not only correlated with reward magnitude for both monetary and erotic rewards, but also with reward prediction error regardless of reward nature (primary or secondary reinforcers) [28] (see Fig. 6.6).

In another fMRI study, we used temporal difference modeling during a classical conditioning learning paradigm. This study investigated prediction error related to different types of reinforcement nature and also compared prediction error for rewards and punishments [29]. Previous fMRI studies using models of reinforcement learning have shown that distinct components of the reward system have a response profile consistent with the temporal difference prediction error signal. However, it was still unclear whether: (1) the reward system discriminates between prediction errors related to reward and punishment or (2) common and distinct brain regions code prediction errors related to different types of outcome. To address these questions, we used a $2 \times 2$ fMRI factorial design crossing the valence (reward and punishment) and the type (taste and vision) of outcomes. Subjects were engaged in a Pavlovian conditioning procedure with four conditions (apple juice, salty water, money, and aversive picture), each with a 50% reinforcement schedule. Trials consisted in two phases: an anticipatory period followed by presentation of the outcome. The results showed that the putamen, the insula, and the anterior cingulate cortex (ACC) code the taste prediction error regardless of valence, that is, respond for both the appetitive and the aversive liquids (juice and salty water). A different pattern of activation was observed in the amygdala, which coded a prediction error only for the primary/immediate reinforcers (apple juice, salty water, and aversive pictures). Finally, the lateral and medial orbitofrontal cortex differentially coded the prediction error for different types of primary reinforcers (liquid versus aversive picture). Indeed, the Blood-oxygen-level (BOLD) activity in the orbitofrontal cortex correlated positively with the prediction error signal for the aversive picture condition and correlated negatively with the prediction error signal for the apple juice and salty water conditions. Taken together, these results demonstrate the different contributions made by distinct brain regions in computing prediction error depending upon the type and valence of the reinforcement (Fig. 6.3).

Finally, a recent approach proposed that temporal-difference signals are not the only learning signals encoded in the brain, in particular when needing to compute the difference between experienced outcomes and outcomes that could have been experienced if decisions had been different (that is a learning signal associated with the actions not taken, i.e., a fictive learning signal) [30]. The authors used a sequential investment task in which after each decision, information was revealed to the subject, regarding whether higher or lower investments would have been a better choice. The natural learning signal for criticizing each choice was the difference between the best return that could have been obtained and the actual gain or loss—that is, the fictive error. Behaviorally, the fictive error was found to be an important determinant for the next investment. The fictive error signal was associated with increasing BOLD response in the ventral caudate nucleus that was not explained by the temporal difference error signal. This fictive error signals may help us to understand the specific roles played by a number of brain regions when making decisions that are subsequently compared to an alternative outcome or decision (counterfactual effect) (see also Chapter 20).

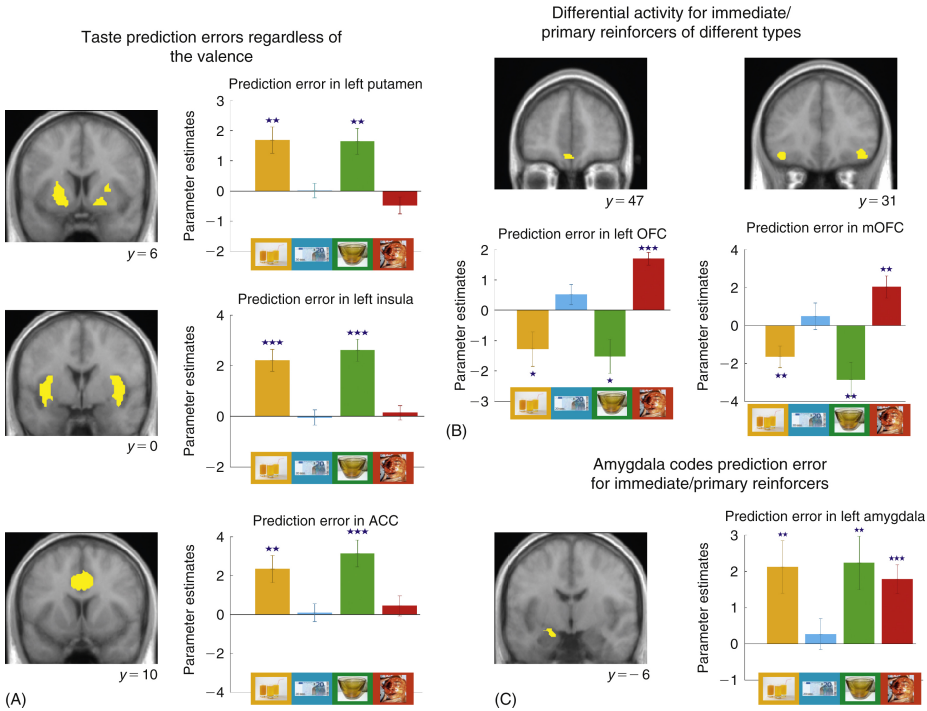


**Figure 6.3** (A) The putamen, the insula, and the ACC code taste prediction error regardless of valence, that is, for both the appetitive and the aversive liquids (juice and salty water). (B) The lateral and medial orbitofrontal cortices differentially code the prediction error for different types of primary reinforcers (liquid versus aversive picture). BOLD activity in the orbitofrontal cortex correlated positively with the prediction error signal for the aversive picture condition and correlated negatively with the prediction error signal for the apple juice and salty water conditions. (C) The amygdala coded a prediction error only for primary/immediate reinforcers (apple juice, salty water, and aversive pictures). Thus, distinct brain regions compute prediction errors depending upon the type and valence of the reinforcement [29].

## 6.3 Computing various uncertainty signals in the brain

We have seen that the prediction error and expected value are crucial signals coded in a number of brain regions, including midbrain dopaminergic neurons and their projection sites. However, recent electrophysiological studies in monkeys indicate that dopaminergic neurons not only code a transient reward prediction error signal but also a sustained signal covarying with reward uncertainty (i.e., reward probability = 0.5) that may be functionally important for risk-seeking behavior and/or exploratory behavior [10]. Until recently, it was unknown whether these two modes of activity could also be observed in humans and whether they could be distinguished by post-synaptic dopaminergic projection sites. Using functional neuroimaging, we have successfully distinguished transient and sustained dynamics of the dopaminergic system in healthy young humans using a new reward task based on the monkey electrophysiology study, which systematically varied monetary

reward probability and magnitude in the absence of choice [20]. The results showed that the human dopaminergic midbrain exhibits similar activity dynamics than midbrain from non-human primates. Moreover, specific dopaminergic projection sites were activated: (a) the ventral striatum, during anticipation of rewards with maximal uncertainty (reward probability = 0.5), (b) the prefrontal cortex and ACCs at the time of the outcome, correlating with a transient prediction error signal coding the difference between expected and obtained rewards (Fig. 6.2). These results indicate that specific functional brain networks subserve the coding of sustained and transient aspects of reward information in humans. These results are important because they support a unified cross-species view in which dopaminergic neurons obey common basic principles of neural computation and provide important new insights into human reward information processing.

It has been proposed that gambling, with its intrinsic reward uncertainty characteristics, has reinforcing properties that may share common mechanisms with addictive drugs [10]. Our results also offers an account for previous reports of human ventral striatum activation during anticipation of monetary and taste rewards for coding, at least in part, the expectation of reward information [22,31,32]. This signal could gain access to striatal neurons through ascending dopamine fibers as well as structures implicated in the evaluation of the motivational significance of stimuli, especially the amygdala and the orbitofrontal cortex.

Our finding of two networks covarying with different reward information signals may indicate that dopaminergic projection sites can distinguish between the two signals. It is also possible that these targets show independent transient (prefrontal cortex) and sustained (ventral striatum) activities related to the two signals and/or that they help to shape dopaminergic neuronal activity by differentially modulating their phasic and sustained modes of firing, which occur independently in individuals neurons [10]. This latter hypothesis is supported by anatomical observations that different populations of dopaminergic neurons are innervated predominantly by the target areas to which they project, or by the regions that, in functional terms, are the most closely linked to the target areas [33]. For example, in rodents, dopaminergic neurons projecting to the prefrontal cortex receive direct reciprocal inputs from this brain region, but not from the striatum, while dopaminergic neurons projecting to the striatum receive afferents from that brain region, but not from the prefrontal cortex, thereby forming two projection systems [33]. This suggests a general principle for midbrain dopaminergic neuronal afferents regulation, the prefrontal cortex and the striatum being responsible for regulating and controlling different modes of dopaminergic neuronal firing.

Another study involving choice behavior investigated the neural correlates of risk, modeled as outcome variance (risk being maximal at 50% probability) found increased activation in the insula, lateral orbitofrontal cortex, and midbrain [24] (see also Chapter 22). Insula activity also correlated with uncertainty in other paradigms involving money and non-monetary stimuli [34,35].

The discrepancy between the different findings of the ventral striatum coding either prediction error or reward uncertainty may be due to several factors. First, most fMRI studies investigating prediction signal used temporal-difference modeling in the context of learning paradigms. In contrast, in our early monetary reward fMRI paradigm [20], there was no learning of cue–outcome associations. So, the putamen activation we observed during anticipation with maximal uncertainty cannot be attributed to a learning effect. Second, one limitation of most fMRI studies varying reward probability is that they could not clearly separate the transient and sustained signals because the delay duration between the conditioned stimulus and the outcome was either too short or randomly

jittered (which is a problem since transient dopaminergic responses are known to depend upon timing uncertainty) [19,24]. To address this problem, we have recently used intracranial recordings in humans to investigate the neural coding of prediction error and uncertainty with a more precise temporal definition (see Fig. 6.4) [36,37].

Although hippocampal–midbrain functional interactions are well documented and the hippocampus receives reward-related information not only from midbrain dopaminergic neurons but also from other components of the reward system, such as the amygdala and orbitofrontal cortex [38], it was still unknown whether it codes statistical properties of reward information, such as prediction error or reward uncertainty. To answer this question, we recorded hippocampal activity in epileptic patients implanted with depth electrodes while they learned to associate cues of slot machines with various monetary reward probabilities (*P*) (unlike our early fMRI monetary reward paradigm in which probability were explicitly given to the subjects) [37] (Fig. 6.4).

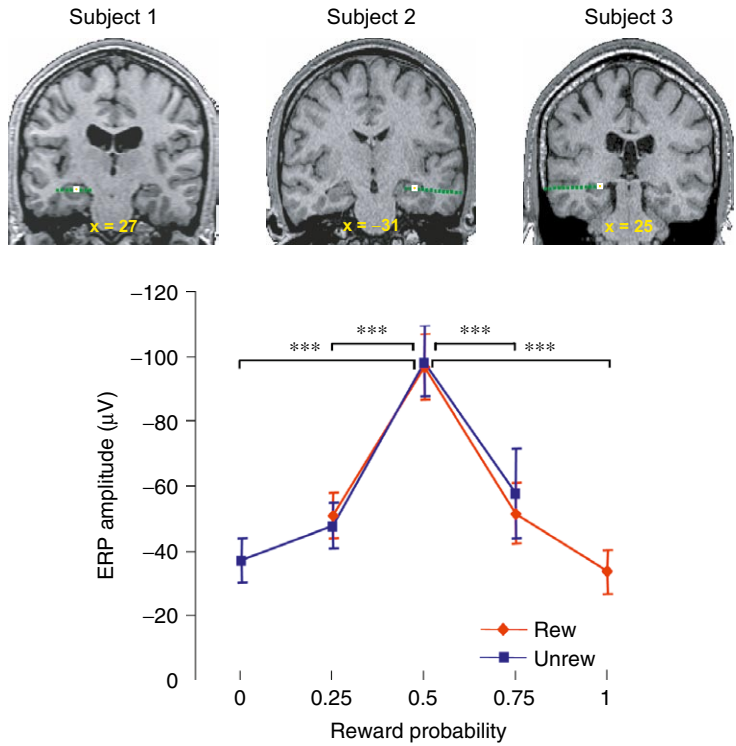


**Figure 6.4** (Top) Location of intracranial electrode contacts. Coronal MRI slices from the three subjects showing the location of the intracranial electrode contacts in the hippocampus. The contacts in the hippocampus yielding the largest potentials are shown in bold square. (Bottom) Uncertainty coding in the human hippocampus. Hippocampal ERP amplitudes code uncertainty at the outcome, regardless of winning or not. Mean peak ERP amplitudes averaged across subjects at the outcome, as a function of reward probability, both for rewarded and for unrewarded trials [37].

Subjects estimated the reward probability of five types of slot machines that varied with respect to monetary reward probabilities $P$ (0 to 1) and that could be discriminated by specific fractal images on top of them. Trials were self-paced and were composed of four distinct phases: (1) Slot machine presentation (S1): subjects pressed one of two response keys to estimate whether the slot machine frequently delivered 20€ or not, based on the outcomes of all the past trials; (2) Delay period (1.5 s): subject's key press triggered three spinners to roll around and to successively stop every 0.5 s during 0.5 s; (3) Outcome S2 (lasting 0.5 €): the third spinner stopped and revealed the trial outcome (i.e., fully informing the subject on subsequent reward or no-reward delivery). Only two configurations were possible at the time the third spinner stopped: "bar, bar, seven" (no reward) or "bar, bar, bar" (rewarded trial); (4) Reward/No-reward delivery (1 s): picture of 20€ bill or rectangle with 0€ written inside.

The results showed that the amplitudes of hippocampal negative event related potentials (ERP), co-varied with uncertainty at the outcome, being maximal for $P = 0.5$ and minimal for $P = 0$ and $P = 1$, regardless of winning or not (Figure 6.4). This inverted U-shape relationship is typical of uncertainty coding and is incompatible with prediction error, novelty, or surprise coding, which would have predicted a negative monotonic correlation between ERP amplitudes and increasing reward probability [10,20]. This uncertainty coding of cue–outcome associations by the hippocampus may constitute a fundamental mechanism underlying the role of this brain region in a number of functions, including attention-based learning, associative learning, probabilistic classification, and binding of stimulus elements, that until now, have received no unified explanation concerning the underlying information processing performed by the hippocampus to achieve them. We propose that the uncertainty coding of cue-outcome associations may constitute the general computational mechanism used by the hippocampus to achieve these different functions. The transient uncertainty signal emitted by the hippocampus at the outcome may play a complementary role to the sustained uncertainty signal emitted by midbrain dopaminergic neurons during the delay period between the cue and the outcome. This finding constitutes a major advance in the knowledge of the functional properties of the human hippocampus and has crucial implications for understanding the basic neural mechanisms used by the brain to extract structural relationships from the environment. It is clear that an ubiquitous coding of uncertainty exists in the human brain, particularly in the midbrain, ventral striatum, insula, ACC, and orbitofrontal cortex [20,24,39–41] and the present study revealed that the hippocampus also participates to uncertainty processing. Future studies are needed to pinpoint the specific roles and time course of each structure in computing uncertainty in different contexts.

## 6.4 Discounting the value of costly options in delay and effort discounting studies

When deciding to engage in a given action, our choice is guided both by the prospect of reward and by the costs that this action entails. Psychological and economic studies have shown that outcome values are discounted with longer delays, an effect known as temporal discounting. A recent electrophysiological study demonstrated that when monkeys choose between sooner smaller available rewards and later larger rewards, the longer the delay of the later larger reward, the less firing of dopaminergic neurons at the time of the conditioned stimuli [8] (see also Chapter 2). Moreover, this reduction in firing rate followed a hyperbolic decay function similar to that observed in choice behavior. In addition, dopamine

responses increased with longer delays at the time of the delayed larger reward delivery, interpreted as reflecting temporal uncertainty and partial learning. These fundamental results establish that dopamine responses reflect the subjective reward value discounted by delay and may provide useful inputs to neural structures involved in intertemporal choices.

Recent fMRI findings on delay-discounting support two opposite theories. According to the first set of experiments, there may be two separate systems in the brain: a limbic system computing the value of rewards delivered immediately or in the near future based on a small discount factor, and a cortical system computing the value of distant rewards based on a high discount factor [17,42–45]. Discounting would result from the interaction of these two systems associated with different value-signals. According to the second theory, based on a recent fMRI study, there would be a single valuation system simply discounting future rewards [46]. One way to conciliate these apparent opposite views is that the striato-prefrontal network might integrate information that is encoded elsewhere in the brain into a single value signal, but that immediate and delayed outcomes activate different types of information that are used to compute the reward value [3]. One further recent finding is that the orbitofrontal cortex may separate the representation of the temporal discount factor applied to distant rewards from the representation of the magnitude of the reward, suggesting that these quantities may be integrated elsewhere in the brain.

Although a few neuroimaging studies start to shed some light on the neural substrates involved in processing subjective value during delay discounting, virtually nothing is known about how effort is discounted in humans. Animal studies demonstrated that the ACC, the ventral striatum, and the orbitofrontal cortex make specific contributions to decision when costly options involve an effort or a delay [47,48]. However, in humans, it is unclear whether there are dissociable pathways underlying different types of costs such as effort and delay to reward. In order to answer this question, we designed a delay/effort discounting task involving primary rewards (visual erotic stimuli) [49]. Heterosexual men were scanned in an event-related fMRI paradigm while performing the task. On every trial, an incentive cue (fuzzy pictures of naked women) briefly appeared on a screen and was followed by the instruction (delay or effort), together with a thermometer indicating the level of delay or effort. Depending on the incentive cue and the proposed cost level, subjects decided whether to invest in the proposed effort (respectively to tolerate the proposed delay) to view the erotic image in clear for 3 s or to perform a minimal effort (respectively to wait for only 1.5 s) to view it for 1 s only. Then, subjects either waited passively in the delay condition (range: 1.5–9 s) or squeezed a hand-grip in the effort condition. We found that choices of the costly option depended upon the subjective value of incentive cues, as indexed by post-scan ratings of these cues, and upon the required level of delay and effort. Thus, decision makers combined two types of information about the benefit (incentive) and cost (level of delay or effort) associated with each option. When investigating the brain regions involved when choosing the costly option regardless of the nature of the cost (delay and effort), we observed stronger activity in the medial anterior prefrontal cortex and in the ventral striatum. These results indicate that choosing the costly option for both types of cost activates common brain regions associated with subjective value coding.

## 6.5 The concept of common neural currency

As noted previously, our behavior is motivated by rewards of different nature among which we frequently need to choose. Because there is no single-sense organ transducing

rewards of different types, our brain must integrate and compare them to choose the options with the highest subjective value. It has been proposed that the brain may use a "common reward currency" that can be used as a common scale to value diverse behavioral acts and sensory stimuli [1]. The need for this common currency arises from the variety of choice we face in our daily life.

Recent behavioral studies in monkeys showed that monkeys differentially value the opportunity to acquire visual information about particular classes of social images. Male rhesus macaques sacrificed fluid for the opportunity to view female perinea and faces of high-status monkeys, but required fluid overpayment to view the faces of low-status monkeys. This work uses a behavioral method to quantify how non-human primates are likely to weigh one type of reward against another [50]. In humans, looking at other people can also be described as rewarding, and that the opportunity to view pictures of the opposite sex is discounted by delay to viewing, substitutes for money, and reinforces work [51]. Attributing value to available options is impaired by orbitofrontal cortex lesion; recent electrophysiological results indicate that some neurons in the orbitofrontal cortex encode the values of offered and chosen goods [52]. Moreover, when a monkey is offered one raisin versus one piece of apple, neurons in the orbitofrontal cortex encode the value of the two goods independently of visuospatial factors and motor responses (contrary to other brain areas in which value modulates activity related to sensory or motor processes). These results make an essential distinction between choosing between goods and choosing between actions. In addition, a classical and general question is how the neuronal representation of value depends upon behavioral context. Although some authors have proposed that the encoded value in the orbitofrontal cortex is relative [53], recent work suggests that neuronal responses in the orbitofrontal cortex are typically invariant for changes of menu—that is, orbitofrontal neuronal response to one particular good usually does not depend on which other goods are available at the same time [54]. These authors proposed that orbitofrontal neuronal activity encodes economic value rather than relative preference.

Because of the properties mentioned previously, the orbitofrontal cortex is likely to be an important brain structure involved in the comparison between different types of goods. However, all the electrophysiological and brain imaging studies published so far compared choices between goods of identical nature (e.g., only food items). Yet, based on the "common currency" concept, there should be a common brain network coding for different types of goods. Many fMRI studies are consistent with this idea, because common brain structures are involved in reward processing, regardless of reward nature. For example, increased midbrain, ventral striatum, and orbitofrontal activities have been observed with different types of rewards, such as monetary gains [19,20,55], pleasant taste [17,18,56], visual erotic stimuli [57,58], beautiful faces [21,59], drugs such as cocaine [60,61] as well as pain relief [62–64]. However, all these neuroimaging studies only investigated one reinforcer at a time and did not compare any two of these reinforcers directly. This was precisely the goal of a recent fMRI study that we performed to compare the common and distinct brain networks involved in processing primary and secondary rewards [28] (see section 6.6).

## 6.6 Common and distinct brain regions involved in processing primary and secondary rewards

Humans are motivated by a wide range of vegetative rewards (such as food and sex) and nonvegetative rewards (such as money, power, fame, and so on). However, it is unclear

whether different types of reinforcers recruit distinct or common neural circuits. In a recent study [28], we compared brain activations to monetary gains and erotic pictures in an incentive delay task. Despite their critical sociobiological importance, visual sexual stimuli have never been studied as reinforcers, but rather as arousing stimuli in passive viewing paradigms focusing on sexual function. They can be considered as "primary rewards," in the sense that they have an innate value and satisfy biological needs. Conversely, money is defined as a "secondary reward," because its value is more abstract and needs to be learned by association with primary rewards.

We hypothesized that monetary and erotic outcomes would activate both shared and distinct cerebral networks. Based on recent fMRI studies, we hypothesized that core components of the reward system, such as the midbrain, ventral striatum, and ACC, would form the core of the shared network ("common currency" network). We also hypothesized a functional dissociation within the orbitofrontal cortex based on a meta-analysis of neuroimaging studies involving different types of rewards. This meta-analysis proposed a postero-anterior dissociation in the orbitofrontal cortex, with more complex or abstract reinforcers being represented more anteriorly than less complex reinforcers [65]. That is, we expected erotic rewards to activate more robustly the posterior part of the orbitofrontal cortex, while the more anterior part of this brain region would be more engaged by secondary rewards. In addition, a crucial question was to know whether the neural correlates of prediction error and expected value could be identified for visual erotic stimuli, which cannot be ascribed an objective value (unlike the amount of monetary reward).

To test our hypotheses, we designed an fMRI experiment comparing brain responses to monetary and visually erotic rewards. Young heterosexual males performed a new event-related fMRI paradigm varying reward nature (money versus erotic stimuli), reward probability, and reward intensity. The structure of each trial was as follows. During anticipation, a cue carried information about the type (monetary or erotic), the probability (0.25, 0.50, or 0.75) and the intensity (high or low) of the upcoming reward. Subjects then had to perform a simple discrimination task by pressing a specified response button for a visual target. At the time of the outcome, they were presented either with "scrambled" pictures (no reward), erotic images, or a picture of a safe indicating an amount of money. At that time, they also had to rate the reward value (of money or erotic stimuli) on a continuous scale.

At the time of outcome, robust BOLD signal was observed for both rewards in a brain circuit including the striatum, the ACC, the midbrain, and the anterior insula. These regions showed a parametric response with the hedonic value, consistent with the idea of a "common neural currency." Moreover, as expected, an antero-posterior dissociation was observed in the lateral orbitofrontal cortex at the time of reward outcome, monetary gains being specifically represented in the anterior part of the orbitofrontal cortex while erotic pictures eliciting activation in its posterior part. This result is important because it identifies a new functional division within the orbitofrontal cortex, with more recent cortical circuits supporting symbolic representation of goods and evolutionary more ancient orbitofrontal region representing subjective value of primary reward (Fig. 6.5).

Another key finding of this study is that prediction error was computed in similar brain regions for monetary and for erotic rewards (Fig. 6.6). Prediction error was defined as the absolute difference between the outcome value and the prediction, where the outcome value was measured by the hedonic ratings and the prediction by the product of expected reward intensity by probability. Brain activity in the ventral striatum, anterior insula, and ACC was shown to positively correlate with prediction error, suggesting that

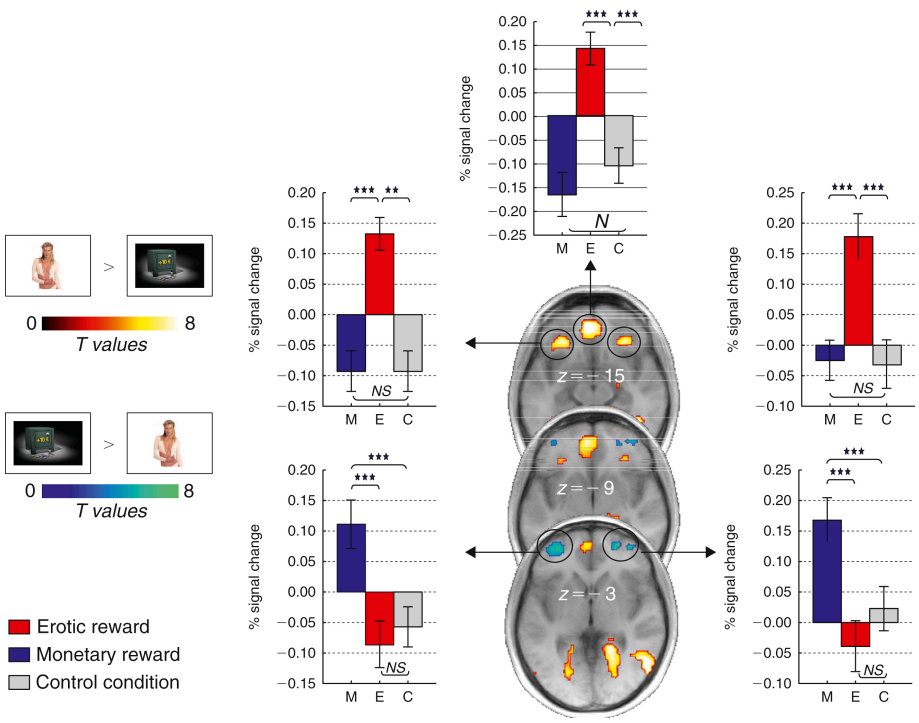


**Figure 6.5** Antero-posterior dissociation within the orbitofrontal cortex according to reward nature. The anterior orbitofrontal cortex codes secondary reward (money) while the posterior and medial orbitofrontal cortex code primary reward (erotic stimuli). Brain regions specifically activated by monetary rewards outcomes are shown in blue-green, and those specifically activated by erotic rewards are shown in red-yellow. Mean percent signal change shows an interaction between reward type and orbitofrontal cortex (OFC) region in both the left and right sides of the brain. Functional maps are overlaid on axial slices of an average anatomical scan of all subjects and are significant at $P < 0.05$ family-wise error (FWE) corrected for multiple comparisons. Asterisks in the bar graphs denote significance of paired comparisons (***$P < 0.001$; **$P < 0.01$; NS, non-significant). Error bars indicate standard error to the mean (SEM) [28]. See Plate 4 of Color Plate section.

prediction error signals might be essentially computed in the brain regions commonly activated by both rewards. These results extend the concept of prediction error to erotic rewards and expand our understanding of reward functions by showing that a common brain network is activated by non-vegetative and vegetative rewards and that distinct orbitofrontal regions respond differentially to various kinds of rewards.

These results are interesting when considering a recent fMRI study suggesting that there may be a single valuation system that discounts future rewards [46]. Another fMRI study supports the idea of a "common neural currency" for two types of rewards [66]. This study showed that the acquisition of one's good reputation robustly activated reward-related brain areas, such as the striatum, and that these areas overlapped with those

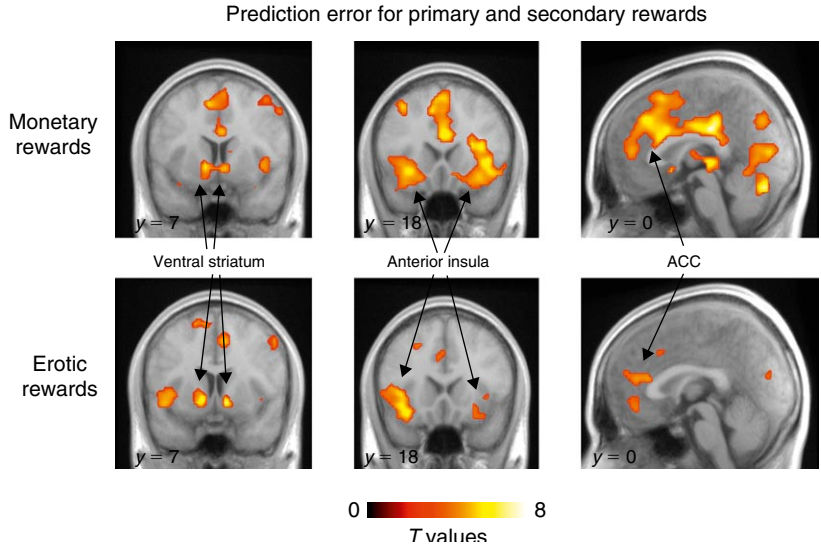


**Figure 6.6** Brain regions responding parametrically with prediction error. Functional maps showing brain regions where BOLD response positively correlates with a measure of prediction error at the time of monetary reward (top) and erotic outcomes (bottom). Prediction error is computed as |Rating – (Probability × Expected intensity)|. Results are overlaid on an average anatomical scan of all subjects and survived a voxel-level threshold of $P < 0.001$ uncorrected for multiple comparisons [28].

activated by monetary rewards. In summary, recent advances in monkey electrophysiology (see Part One of this book) and in functional neuroimaging suggest that individuals use some of the same circuits to process money and other types of rewards, even in the absence of choice between them.

## 6.7 Distinguishing two brain systems involved in choosing between different types of rewards

Because the concept of "common currency" involves the notion of comparison between different types of goods, we have also recently characterized distinct value-based signals involved when choosing between different types of rewards [67]. One signal reflects the computation of the distance between the subjective values of each option. Another one codes the subjective value of the chosen option, while a third elementary signal involved in motivated choice codes choice uncertainty (maximal at the point of subjective equivalence between the two options).

Young heterosexual males, drink-deprived for 12 h, were scanned in a new fMRI paradigm while choosing between two gambles, one rewarded by a very small amount of fruit juice (0.5 ml) and the other by visual erotic stimuli (pictures of naked women) (Fig. 6.1A). Participants experienced both types of primary rewards directly inside the scanner. For each trial, two pie charts indicated the reward probabilities, varying independently (e.g., $P = 0.75$ juice versus $P = 0.5$ erotic stimulus) (Fig. 6.7.A). One important aspect of the task is that

the magnitude of the reward was kept constant. Therefore, choices were made on the basis of preference for a type of reward and on the basis of reward probability.

We first estimated the preference of each participant for fruit juice over an erotic picture and expressed it as an equivalent offer by fitting, for each participant, a logistic model of the probability of choice that included the probability of being rewarded by the fruit juice, the erotic picture, and the trial number as explanatory variables. This last variable accounted for a possible drift of the preference during the experiment and was included in the model as a control. The preference was computed as the ratio of the parameter estimates for the picture and drink. Then, the subjective distance between options for each offer was computed as the difference between the subjective value of the juice option and the subjective value of the erotic picture option.

Behavioral results indicated that participants had heterogeneous preferences, some preferring juice over pictures, others pictures over juice. Response times increased linearly with choice uncertainty, indicating that the decision process slows down as the subjective distance between options decreases and as it becomes harder to discriminate which option is the best. Conversely, response times decreased as the subjective value of the chosen option increased, reflecting higher motivation for the favored choice. Moreover, the proportion of choice of a given option varied as a sigmoidal function of the distance between the subjective values of each option, showing that probability of choice is effectively modulated by the difference between subjective values of the available options (Fig. 6.7.B).

The brain imaging results revealed that, with increasing difference between subjective values, activity increased in the medial anterior and lateral parts of the orbitofrontal cortex and the midbrain, reflecting computation of the distance between options in a "common currency" space (Fig. 6.7.D). The same orbitofrontal regions coding the subjective distance between options at the time of decision also coded the subjective value of the chosen option. At the time of the rewarded outcome, the error prediction signal varied in the same direction as the subjective difference between options: it decreased when juice was delivered as compared to when it was not delivered, and it increased when a picture was delivered relative to when it was not (Fig. 6.7.E).

Moreover, brain regions coding choice uncertainty involved the ACC, the bilateral anterior insula and the inferior frontal gyri. This activity is likely to reflect the slowing down of the decision process observed behaviorally. Importantly, BOLD activity in the orbitofrontal cortex did not correlate with choice uncertainty, even when lowering the statistical threshold. Overall, these results indicate a functional dissociation between two brain networks: the orbitofrontal cortex, which codes the subjective values related to the goal of the decision, and the ACC/anterior insula network, which codes the uncertainty on these values.

Moreover, these results indicate that the same orbitofrontal cortex region codes different value-related signals and pinpoint a brain network composed of the ACC and the anterior insula that computes choice uncertainty.

## 6.8 Goal-value and decision-value signals

It is still unclear how many value-related signals are computed by the brain to make decisions. The computational complexity of these signals, as well as their possible reducibility to canonical signals remain also poorly characterized. Two value signals that may be computed during decision making are goal-value and decision-value signals, which may be used to choose the option with highest benefit. Goal values measure the predicted amount of reward associated with the outcome generated by each of the actions under consideration. Decision values measure the net value of taking the different actions, that

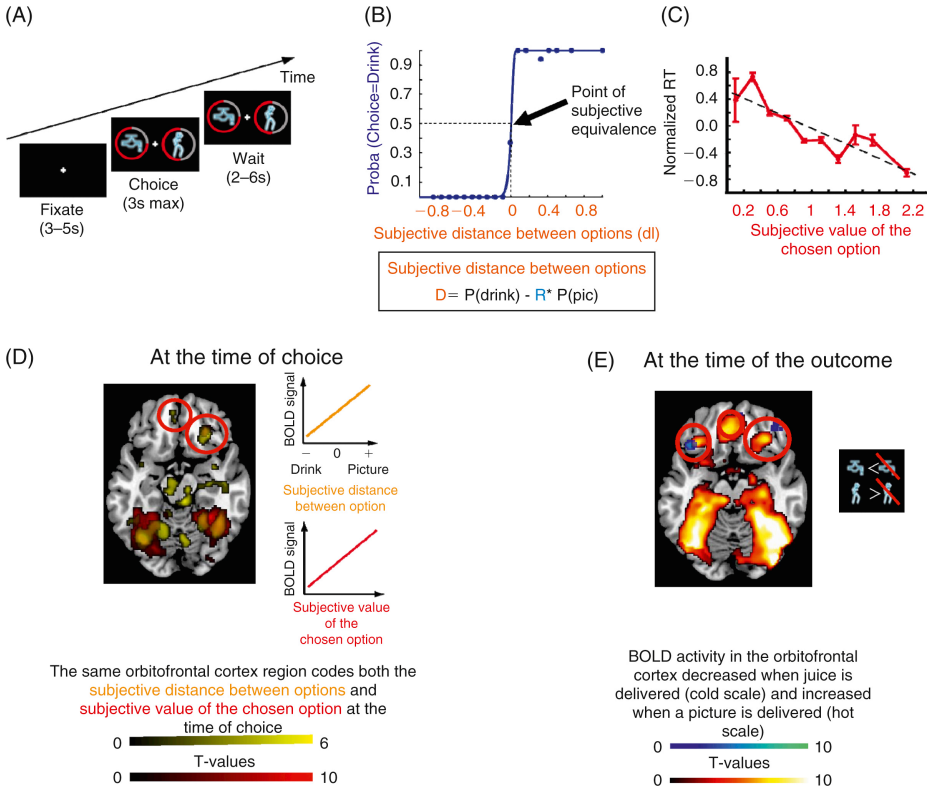


**Figure 6.7** (A) Design of the experiment. Participants were asked to freely make motivated choices between two types of rewards with varying reward probability. For example, the offer consisted in a choice between an option rewarded 75% of the time by fruit juice and by an option rewarded 50% of the time by an erotic picture. The red part of the circle around each symbol indicates the reward probability. After a waiting period lasting 2-6 seconds, participants were probabilistically rewarded with juice or by viewing an erotic picture. (B) Estimation of the preference of each participant for drinking fruit juice over viewing erotic pictures was expressed as an equivalent offer by fitting a logistic model of the probability of choice that included the probability of being rewarded by a drink, by a picture and by the trial number. The preference was computed as the ratio of the betas for the picture and the drink. The subjective distance between options was computed for each offer as the difference between the subjective value of the drink option and the subjective value of the picture option. (C) Response times decreased as the subjective value of the chosen option increased. (D) At the time of choice, the same orbitofrontal regions coding the subjective distance (yellow scale) between options also coded the subjective value of the chosen option (red). (E) At the time of the outcome, the error prediction signal varied in the same direction as the subjective difference between options: it decreased at the time of reward when juice was delivered compared to when it was not delivered (cold scale), and it increased when a picture was delivered compared to when is was not viewed (hot scale). See plate 5 of color plate section.

is, the benefits *minus* the costs. In a recent fMRI study, goal value was computed by the willingness to pay for different food items, while decision values were computed by subtracting the price of the offered food items from the goal value [68]. These authors found that activity in the medial orbitofrontal cortex were correlated with goal values, that activity in the central orbitofrontal cortex correlated with decision values, and that activity in the ventral striatum correlated with prediction errors.

To conclude, the studies reviewed here indicate that the human orbitofrontal cortex is involved in processing a number of value signals, such as the subjective values of stimuli, but also contributes to processing signals related to the decision-making process itself, such as the distance between the subjective value of different options or the subjective distance of chosen option, thereby coding signals informing about what action to take next.

## 6.9 Variation in dopamine genes influence reward processing

Both reward processing and decision making engage brain structures that lie on the ascending dopaminergic pathways. An important axis of current research is to study the brain influence of genes that affect dopaminergic transmission in order to clarify the biological mechanisms underlying interindividual differences and vulnerability to pathology related to the dysfunction of the dopaminergic system. Although there are clear individual genetic differences regarding susceptibility to and manifestation of these neuropsychopathologies, the influence of genetic predispositions and variations on activation of the human reward system remains poorly understood. Recent neuroimaging and behavioral studies have focused on the genetic variations of dopamine receptors, especially DRD2 and DRD4, and other genes coding for enzymes and transporters involved in the dopaminergic transmission, such as the catechol-*O*-methyltransferase (COMT) and the dopamine transporter (DAT). For example, polymorphisms in dopamine receptor (DRD4) and monoamine oxidase A (MAOA) genes showed significant associations with efficiency of handling conflict as measured by reaction time differences in an attention task and modulate ACC activation [69]. Moreover, the role of the DRD2 polymorphism in monitoring negative action outcomes and feedback-based learning was tested during a probabilistic learning task [70]. A1-allele carriers, with reduced dopamine D2 receptor densities, showed lower posterior medial frontal cortex activity, involved in feedback monitoring, and learned to avoid actions with negative consequences less efficiently. The authors suggested that dopamine D2 receptor reduction seems to decrease sensitivity to negative action consequences, which may explain an increased risk of developing addictive behaviors in A1-allele carriers. Recent behavioral and computational modeling works also suggest independent gene effects (DARPP-32, DRD2, COMT) on reinforcement learning parameters that contribute to reward and avoidance learning in humans. These findings support a neurocomputational dissociation between striatal and prefrontal dopaminergic mechanisms in reinforcement learning [71] (see also Chapter 19).

Two important proteins contribute to terminating the action of intrasynaptic dopamine in the brain: COMT, which catabolizes released dopamine, and the DAT, which plays a crucial role in determining the duration and amplitude of dopamine action by rapidly recapturing extracellular dopamine into presynaptic terminals after release. In humans, the *COMT* gene contains a common and evolutionarily recent functional polymorphism that codes for the substitution of valine (val) by methionine (met) at codon 158, referred to as $Val^{158}Met$ polymorphism. The COMT enzyme is involved in the metabolic degradation of catecholamines, converting dopamine into 3-methoxytyramine and norepinephrine into normetanephrine. Because the COMT protein containing methionine is

relatively thermolabile, its activity is lower at body temperatures than the COMT valine protein, which is fully active at body temperature. Hence, individuals with two copies of the met allele (met/met) have 25–75% reduction in COMT enzyme activity, and therefore presumptively more baseline synaptic dopamine, compared to individuals with two copies of the val allele (val/val) [72,73].

The *DAT*1 gene (SLC6A3) includes 15 exons, with a variable number of tandem repeat (VNTR) polymorphisms in the 15th exon, a region encoding the transcript's 3′ UTR [74]. The 40-bp VNTR element is repeated between 3 and 13 times but in most of the population occurs with greatest frequency in the 9- and 10-repeat forms. The expression of the *DAT*1 9-repeat allele is lower than the 10-repeat allele [75–77], although one study reported the opposite allelic associations [78]. Thus, the *DAT*1 10-repeat allele, associated with increased expression of the gene, presumably leads to relatively decreased extrasynaptic striatal dopamine levels. This is consistent with a human SPECT study reporting increased striatal DAT availability in 9-repeat carriers relative to 10-repeat carriers [79], although another study failed to support this [75]. Mice lacking the *DAT*1 gene show extensive adaptative changes in the dopaminergic system, the DAT controlling both the duration of extracellular dopamine signals and regulating presynaptic dopamine homeostasis [80].

Importantly, animal studies indicate differential functional localization of the COMT and DAT proteins. The COMT enzyme plays a particular role in modulating dopamine in the prefrontal cortex, where *DAT*1 expression is sparse [81,82]. *COMT* is expressed more abundantly in cortical neurons than in the striatum [83], but it is unclear to what extent *COMT* modulates catecholamine function outside the cortex. Recent studies in COMT knockout mice suggest that COMT has little if any role in striatal DA levels [84]. In contrast, animal research and human postmortem studies indicate that the *DAT*1 is expressed abundantly in midbrain, striatum, and hippocampus but sparsely in the prefrontal cortex (PFC) [85,86].

In parallel with the fundamental fMRI results concerning prediction error mentioned before, fMRI studies in healthy young subjects have documented that distinct reward anticipation- and outcome-processing phases are associated with differential patterns of specific midbrain dopaminergic postsynaptic targets [20,32,87]. Specifically, anticipation of reward robustly activates foci in the ventral striatum [32,87], particularly during anticipation of rewards with maximal uncertainty (i.e., reward probability = 0.5) [20] while rewarded outcomes activate the lateral and orbital parts of the PFC [20,87]. Despite the direct involvement of the COMT and DAT proteins in dopamine transmission, the influences of *COMT* and *DAT*1 functional polymorphisms on distinct components of the reward system have not been as systematically explored as have been the domains of working and episodic memory [86,88,89].

We recently used event-related fMRI and a recently developed reward paradigm to directly investigate the relationship between *COMT* and *DAT*1 functional polymorphisms and the response of the reward system during anticipation of uncertain rewards and at the time of reward delivery, bridging the gap between basic molecular genetics, fundamental electrophysiological findings, and functional neuroimaging in humans [90].

The results revealed a main effect of *COMT* genotype in the ventral striatum and lateral prefrontal cortex during reward anticipation and in the orbitofrontal cortex at the time of reward delivery, met/met individuals exhibiting the highest activation (Fig. 6.8).

The main effect of *DAT*1 genotype was seen in robust BOLD response differences in the caudate nucleus and ventral striatum during reward anticipation and in the lateral prefrontal cortex and midbrain at the time of reward delivery, with carriers of the *DAT*1 9-repeat allele showing the highest activity. Moreover, an interaction between the *COMT* and *DAT*1 genes was found in the ventral striatum and lateral prefrontal cortex during

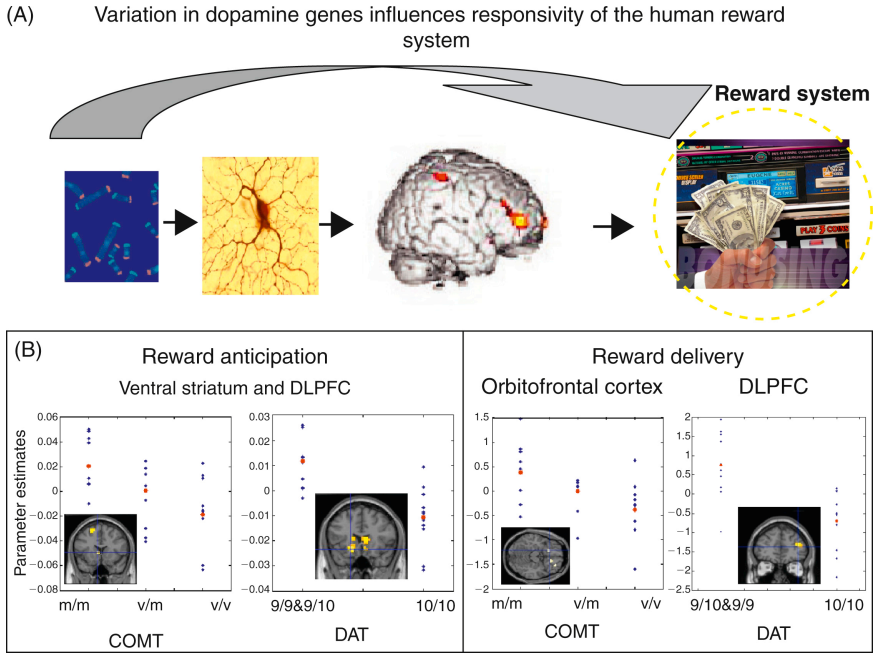


**Figure 6.8** (A) Relationships between the effects of genetic variations and reward processing. Influence of the polymorphisms of the catecholamine-O-methyltransferase (COMT) (valine/valine; valine/methionine; methionine/methionine) and the dopamine transporter (DAT) (9/9&9/10; 10/10) on the reward system. (B) (Left) Main effect of *COMT* and *DAT* genotypes during anticipation of reward with maximal uncertainty. Negative relationship was observed between *COMT* val allele dosage (0_met/met, 1_val/met, or 2_val/val) and BOLD response in the ventral striatum, left superior PFC, and dorsolateral PFC during anticipation of reward with maximal uncertainty. More robust BOLD response was observed in 9-repeat carriers (including *DAT*1 9-repeat and 9/10) compared to 10-repeat individuals during reward anticipation in the bilateral ventral striatum. (Right) Main effect of *COMT* and *DAT* genotypes at the time of reward delivery. Negative relationship between *COMT* val allele dosage and orbitofrontal cortex activation at the time of reward delivery. Higher lateral prefrontal BOLD signal was observed in *DAT*1 9-repeat allele dosage compared to 10-repeat carriers at the time of reward delivery [90].

reward anticipation and in the lateral prefrontal and orbitofrontal cortices as well as in the midbrain at the time of reward delivery, with carriers of the *DAT*1 9-repeat allele and *COMT* met/met allele exhibiting the highest activation, presumably reflecting functional change consequent to higher synaptic dopamine availability.

One important insight provided by our data is a clear demonstration of interaction between the *DAT*1 and *COMT* genes that control a complex phenotype (activation of the reward system). This interaction likely reflects differences in dopaminergic level due to the combined effect of the *COMT* val/val and *DAT*1 10/10 alleles on elimination of DA in the fronto-striatal system. Interestingly, the effects on the BOLD signal of this presumed low DA level in val/val and 10-repeat alleles' carriers differ both according to brain regions and task phases.

These results indicate that genetically-influenced variations in dopamine transmission modulate the response of brain regions involved in anticipation and reception of rewards and suggest that these responses may contribute to individual differences in reward-seeking behavior and in predisposition to neuropsychiatric disorders.

A recent study used a guessing task to investigate how individual variation in *COMT* and *DAT*1 genes influences reward processing [91] (see also Chapters 16 and 17). In accordance with our results, this study reported that, during reward anticipation, the lateral PFC and the ventral striatum activities were *COMT* genotype-dependent: subjects homozygous for the met allele showed larger responses in these brain regions compared with volunteers homozygous for the val allele. This effect was observed when averaging all probabilities and magnitudes against baseline, but no main effect of *COMT* genotype was observed on ventral striatal sensitivity to reward uncertainty. Moreover, no main effect of *DAT*1 genotype was reported on striatal activity during reward anticipation, despite the well-established abundancy of DAT in the striatum. A gene–gene interaction between *COMT* and *DAT*1 was observed in the ventral striatum when sorting genotypes from met/met *DAT*1 10-repeat allele to val/val 9-repeat allele, interpreted as consistent with the notion that basal dopaminergic tone, regulated by *COMT*, interacts with phasic dopamine release, regulated by the *DAT*. It is difficult to directly compare our findings to these results because *COMT* and *DAT*1 genotypes may both directly influence distinct components of the human reward system (COMT modulating the dorsolateral prefrontal cortex (DLPFC) and DAT the striatum) and differentially affect their neurofunctional balance in a task-dependent manner. Finally, because this previous study did not report effects of genotype on fMRI results at the time of reward delivery, it remains unclear whether distinct phases of this guessing task induce differential brain activity dependent upon *COMT* and *DAT*1 polymorphisms.

It should be noted that our fMRI results on COMT/DAT genotypes cannot establish the neurophysiological mechanisms underlying the relationship between dopamine release and BOLD signal increase [92]. However, our study directly links genotype-dependent synaptic dopamine availability with BOLD signal change in humans and suggests that higher BOLD signal at prefronto-striatal sites is associated with greater dopamine synaptic availability (i.e., lower DA elimination), in agreement with recent studies observing: (a) that in young adults there is a tight coupling between increased midbrain dopamine synthesis and reward-related increase BOLD signal in the PFC both during reward anticipation and at the time of reward delivery [93]; and (b) that in animals, injection of dopamine-releasing agents increases BOLD signal in mesolimbic regions (frontal cortex, striatum, cingulate cortex) with a time course that parallels the changes observed by microdialysis measurements of striatal dopamine release [94].

## 6.10 Conclusions

Making choices requires processing of several value-related signals, such as prediction error, uncertainty, subjective value of different options and the distance between them, goal value, and decision value. In this review, we have described neuroimaging evidence showing the neural substrates of these different value signals. The integrity of the neural structures computing these value signals are crucial for efficient decision making and processing of reward information. Clinical areas of research in which the current knowledge on value-based decision making can be applied concern a variety of neuropathologies, such as schizophrenia, Parkinson's disease, pathological gambling, or drug

addiction. A better knowledge of the neural basis of goal values, decision values, and prediction errors is likely to advance our understanding of the impact that different types of neuropathologies have on reward and decision making. For example, a recent neuroimaging study relates dysfunctions in motivational salience and prediction error signal to explain the abnormal mental experience of psychosis. Patients with psychosis exhibited abnormal physiological responses associated with reward prediction error in the dopaminergic midbrain, striatum, and limbic system, providing the first evidence linking abnormal mesolimbic activity, reward learning and psychosis [95] (see also Chapter 11). Future works should also investigate how genetically-influenced variations in different monoamine transmitters (dopamine, noradrenaline, serotonin), modulate the response of brain regions coding the various value signals involved in decision making outlined in this chapter.

## Acknowledgments

I thank Philippe Domenech, Elise Météreau, Charlotte Prévost, Guillaume Sescousse and Dr Giovanna Vanni-Mercier for performing some of the experiments described here and for their helpful comments on early versions of this chapter. J-C Dreher is supported by the CNRS, the Fyssen foundation and by the "Fondation pour la Recherche Médicale".

## References

[1] L.P. Sugrue, G.S. Corrado, W.T. Newsome, Choosing the greater of two goods: neural currencies for valuation and decision making, Nat. Rev. Neurosci. 6 (2005) 363–375.
[2] K. Doya, Modulators of decision making, Nat. Neurosci. 11 (2008) 410–416.
[3] A. Rangel, C. Camerer, P.R. Montague, A framework for studying the neurobiology of value-based decision making, Nat. Rev. Neurosci. 9 (2008) 545–556.
[4] W. Schultz, Multiple reward signals in the brain, Nat. Rev. Neurosci. 1 (2000) 199–207.
[5] W. Schultz, A. Dickinson, Neuronal coding of prediction errors, Annu. Rev. Neurosci. 23 (2000) 473–500.
[6] H.M. Bayer, B. Lau, P.W. Glimcher, Statistics of midbrain dopamine neuron spike trains in the awake primate, J. Neurophysiol. 98 (2007) 1428–1439.
[7] P.N. Tobler, C.D. Fiorillo, W. Schultz, Adaptive coding of reward value by dopamine neurons, Science 307 (2005) 1642–1645.
[8] S. Kobayashi, W. Schultz, Influence of reward delays on responses of dopamine neurons, J. Neurosci. 28 (2008) 7837–7846.
[9] M.R. Roesch, D.J. Calu, G. Schoenbaum, Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards, Nat. Neurosci. 10 (2007) 1615–1624.
[10] C.D. Fiorillo, P.N. Tobler, W. Schultz, Discrete coding of reward probability and uncertainty by dopamine neurons, Science 299 (2003) 1898–1902.
[11] G. Morris, A. Nevet, D. Arkadir, E. Vaadia, H. Bergman, Midbrain dopamine neurons encode decisions for future action, Nat. Neurosci. 9 (2006) 1057–1063.
[12] H. Ji, P.D. Shepard, Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism, J. Neurosci. 27 (2007) 6923–6930.
[13] M. Matsumoto, O. Hikosaka, Lateral habenula as a source of negative reward signals in dopamine neurons, Nature 447 (2007) 1111–1115.

[14] S. Hong, O. Hikosaka, The globus pallidus sends reward-related signals to the lateral habenula, Neuron 60 (2008) 720–729.
[15] G.S. Berns, S.M. McClure, G. Pagnoni, P.R. Montague, Predictability modulates human brain response to reward, J. Neurosci. 21 (2001) 2793–2798.
[16] P.C. Fletcher, J.M. Anderson, D.R. Shanks, R. Honey, T.A. Carpenter, T. Donovan, N. Papadakis, E.T. Bullmore, Responses of human frontal cortex to surprising events are predicted by formal associative learning theory, Nat. Neurosci. 4 (2001) 1043–1048.
[17] S.M. McClure, G.S. Berns, P.R. Montague, Temporal prediction errors in a passive learning task activate human striatum, Neuron 38 (2003) 339–346.
[18] J.P. O'Doherty, P. Dayan, K. Friston, H. Critchley, R.J. Dolan, Temporal difference models and reward-related learning in the human brain, Neuron 38 (2003) 329–337.
[19] B. Abler, H. Walter, S. Erk, H. Kammerer, M. Spitzer, Prediction error as a linear function of reward probability is coded in human nucleus accumbens, Neuroimage 31 (2006) 790–795.
[20] J.C. Dreher, P. Kohn, K.F. Berman, Neural coding of distinct statistical properties of reward information in humans, Cereb. Cortex 16 (2006) 561–573.
[21] S. Bray, J. O'Doherty, Neural coding of reward-prediction error signals during classical conditioning with attractive faces, J. Neurophysiol. 97 (2007) 3036–3045.
[22] B. Knutson, C.M. Adams, G.W. Fong, D. Hommer, Anticipation of increasing monetary reward selectively recruits nucleus accumbens, J. Neurosci. 21 (2001) RC159.
[23] B. Knutson, J. Taylor, M. Kaufman, R. Peterson, G. Glover, Distributed neural representation of expected value, J. Neurosci. 25 (2005) 4806–4812.
[24] K. Preuschoff, P. Bossaerts, S.R. Quartz, Neural differentiation of expected reward and risk in human subcortical structures, Neuron 51 (2006) 381–390.
[25] J. Yacubian, J. Glascher, K. Schroeder, T. Sommer, D.F. Braus, C. Buchel, Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain, J. Neurosci. 26 (2006) 9530–9537.
[26] S. Kobayashi, J. Lauwereyns, M. Koizumi, M. Sakagami, O. Hikosaka, Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex, J. Neurophysiol. 87 (2002) 1488–1498.
[27] M. Watanabe, K. Hikosaka, M. Sakagami, S. Shirakawa, Coding and monitoring of motivational context in the primate prefrontal cortex, J. Neurosci. 22 (2002) 2391–2400.
[28] G. Sescousse, J-C. Dreher, Coding of reward type along an antero-posterior gradient in the human orbitofrontal cortex, in: Exciting Biologies meeting, Chantilly, France, 2008. Biology of Cognition.
[29] E. Météreau, J-C. Dreher, Neural responses underlying predictive learning of different types of rewards and punishments, in: Society for Neuroscience, Washington DC, 2008.
[30] P.H. Chiu, T.M. Lohrenz, P.R. Montague, Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task, Nat. Neurosci. 11 (2008) 514–520.
[31] H.C. Breiter, I. Aharon, D. Kahneman, A. Dale, P. Shizgal, Functional imaging of neural responses to expectancy and experience of monetary gains and losses, Neuron 30 (2001) 619–639.
[32] J.P. O'Doherty, R. Deichmann, H.D. Critchley, R.J. Dolan, Neural responses during anticipation of a primary taste reward, Neuron 33 (2002) 815–826.
[33] S.R. Sesack, D.B. Carr, N. Omelchenko, A. Pinto, Anatomical substrates for glutamate-dopamine interactions: evidence for specificity of connections and extrasynaptic actions, Ann. NY Acad. Sci. 1003 (2003) 36–52.
[34] S.A. Huettel, A.W. Song, G. McCarthy, Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices, J. Neurosci. 25 (2005) 3304–3311.
[35] J. Grinband, J. Hirsch, V.P. Ferrera, A neural representation of categorization uncertainty in the human brain, Neuron 49 (2006) 757–763.

[36] J. Thomas, G. Vanni-Mercier, J-C. Dreher, Temporal dynamics of reward probability coding: a Magnetoencephalographic study in humans, in: Human Brain Mapping meeting, Melbourne, Australia, 2008.
[37] G. Vanni-Mercier, F. Mauguière, J. Isnard, J-C. Dreher, The hippocampus codes the uncertainty of cue-outcome associations: an intracranial electrophysiological study in humans. Journal of Neuroscience 29: 5287–5294.
[38] W.A. Suzuki, D.G. Amaral, Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices, J. Neurosci. 14 (1994) 1856–1877.
[39] M. Hsu, M. Bhatt, R. Adolphs, D. Tranel, C.F. Camerer, Neural systems responding to degrees of uncertainty in human decision-making, Science 310 (2005) 1680–1683.
[40] P.N. Tobler, J.P. O'Doherty, R.J. Dolan, W. Schultz, Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems, J. Neurophysiol. 97 (2007) 1621–1632.
[41] K. Preuschoff, S.R. Quartz, P. Bossaerts, Human insula activation reflects risk prediction errors as well as risk, J. Neurosci. 28 (2008) 2745–2752.
[42] S.C. Tanaka, K. Doya, G. Okada, K. Ueda, Y. Okamoto, S. Yamawaki, Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops, Nat. Neurosci. 7 (2004) 887–893.
[43] S.M. McClure, K.M. Ericson, D.I. Laibson, G. Loewenstein, J.D. Cohen, Time discounting for primary rewards, J. Neurosci. 27 (2007) 5796–5804.
[44] N. Schweighofer, S.C. Tanaka, K. Doya, Serotonin and the evaluation of future rewards: theory, experiments, and possible neural mechanisms, Ann. NY Acad. Sci. 1104 (2007) 289–300.
[45] N. Schweighofer, M. Bertin, K. Shishida, Y. Okamoto, S.C. Tanaka, S. Yamawaki, K. Doya, Low-serotonin levels increase delayed reward discounting in humans, J. Neurosci. 28 (2008) 4528–4532.
[46] J.W. Kable, P.W. Glimcher, The neural correlates of subjective value during intertemporal choice, Nat. Neurosci. 10 (2007) 1625–1633.
[47] M.E. Walton, S.W. Kennerley, D.M. Bannerman, P.E. Phillips, M.F. Rushworth, Weighing up the benefits of work: behavioural and neural analyses of effort-related decision making, Neural Netw. 19 (2006) 1302–1314.
[48] M.F. Rushworth, T.E. Behrens, P.H. Rudebeck, M.E. Walton, Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour, Trends Cogn. Sci. 11 (2007) 168–176.
[49] C. Prévost, M. Pessiglione, M-C. Cléry-Melin, J-C. Dreher, Delay versus effort-discounting in the human brain, in: Society for Neuroscience, Washington DC, 2008.
[50] R.O. Deaner, A.V. Khera, M.L. Platt, Monkeys pay per view: adaptive valuation of social images by rhesus macaques, Curr. Biol. 15 (2005) 543–548.
[51] B.Y. Hayden, M.L. Platt, Temporal discounting predicts risk sensitivity in rhesus macaques, Curr. Biol. 17 (2007) 49–53.
[52] C. Padoa-Schioppa, J.A. Assad, Neurons in the orbitofrontal cortex encode economic value, Nature 441 (2006) 223–226.
[53] L. Tremblay, W. Schultz, Relative reward preference in primate orbitofrontal cortex, Nature 398 (1999) 704–708.
[54] C. Padoa-Schioppa, J.A. Assad, The representation of economic value in the orbitofrontal cortex is invariant for changes of menu, Nat. Neurosci. 11 (2008) 95–102.
[55] J.P. O'Doherty, Reward representations and reward-related learning in the human brain: insights from neuroimaging, Curr. Opin. Neurobiol. 14 (2004) 769–776.
[56] J. O'Doherty, Can't learn without you: predictive value coding in orbitofrontal cortex requires the basolateral amygdala, Neuron 39 (2003) 731–733.

[57] J. Redoute, S. Stoleru, M.C. Gregoire, N. Costes, L. Cinotti, F. Lavenne, D. Le Bars, M.G. Forest, J.F. Pujol, Brain processing of visual sexual stimuli in human males, Hum. Brain Mapp. 11 (2000) 162–177.

[58] S. Karama, A.R. Lecours, J.M. Leroux, P. Bourgouin, G. Beaudoin, S. Joubert, M. Beauregard, Areas of brain activation in males and females during viewing of erotic film excerpts, Hum. Brain Mapp. 16 (2002) 1–13.

[59] J.S. Winston, J. O'Doherty, J.M. Kilner, D.I. Perrett, R.J. Dolan, Brain systems for assessing facial attractiveness, Neuropsychologia 45 (2007) 195–206.

[60] R.C. Risinger, B.J. Salmeron, T.J. Ross, S.L. Amen, M. Sanfilipo, R.G. Hoffmann, A.S. Bloom, H. Garavan, E.A. Stein, Neural correlates of high and craving during cocaine self-administration using BOLD fMRI, Neuroimage 26 (2005) 1097–1108.

[61] P. Kufahl, Z. Li, R. Risinger, C. Rainey, L. Piacentine, G. Wu, A. Bloom, Z. Yang, S.J. Li, Expectation modulates human brain responses to acute cocaine: a functional magnetic resonance imaging study, Biol. Psychiatry 63 (2008) 222–230.

[62] B. Seymour, J.P. O'Doherty, P. Dayan, M. Koltzenburg, A.K. Jones, R.J. Dolan, K.J. Friston, R. S. Frackowiak, Temporal difference models describe higher-order learning in humans, Nature 429 (2004) 664–667.

[63] B. Seymour, J.P. O'Doherty, M. Koltzenburg, K. Wiech, R. Frackowiak, K. Friston, R. Dolan, Opponent appetitive-aversive neural processes underlie predictive learning of pain relief, Nat. Neurosci. 8 (2005) 1234–1240.

[64] B. Seymour, T. Singer, R. Dolan, The neurobiology of punishment, Nat. Rev. Neurosci. 8 (2007) 300–311.

[65] M.L. Kringelbach, The human orbitofrontal cortex: linking reward to hedonic experience, Nat. Rev. Neurosci. 6 (2005) 691–702.

[66] K. Izuma, D.N. Saito, N. Sadato, Processing of social and monetary rewards in the human striatum, Neuron 58 (2008) 284–294.

[67] P. Domenech, J-C. Dreher, Distinguishing two brain systems involved in choosing between different types of rewards, in: Society for Neuroscience, Washington DC, 2008.

[68] T.A. Hare, J. O'Doherty, C.F. Camerer, W. Schultz, A. Rangel, Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors, J. Neurosci. 28 (2008) 5623–5630.

[69] J. Fan, J. Fossella, T. Sommer, Y. Wu, M.I. Posner, Mapping the genetic variation of executive attention onto brain activity, Proc. Natl. Acad. Sci. USA 100 (2003) 7406–7411.

[70] T.A. Klein, J. Neumann, M. Reuter, J. Hennig, D.Y. von Cramon, M. Ullsperger, Genetically determined differences in learning from errors, Science 318 (2007) 1642–1645.

[71] M.J. Frank, A.A. Moustafa, H.M. Haughey, T. Curran, K.E. Hutchison, Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning, Proc. Natl. Acad. Sci. USA 104 (2007) 16311–16316.

[72] H.M. Lachman, D.F. Papolos, T. Saito, Y.M. Yu, C.L. Szumlanski, R.M. Weinshilboum, Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders, Pharmacogenetics 6 (1996) 243–250.

[73] J. Chen, B.K. Lipska, N. Halim, Q.D. Ma, M. Matsumoto, S. Melhem, B.S. Kolachana, T. M. Hyde, M.M. Herman, J. Apud, M.F. Egan, J.E. Kleinman, D.R. Weinberger, Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain, Am. J. Hum. Genet. 75 (2004) 807–821.

[74] D.J. Vandenbergh, A.M. Persico, A.L. Hawkins, C.A. Griffin, X. Li, E.W. Jabs, G.R. Uhl, Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR, Genomics 14 (1992) 1104–1106.

[75] A. Heinz, D. Goldman, D.W. Jones, R. Palmour, D. Hommer, J.G. Gorey, K.S. Lee, M. Linnoila, D.R. Weinberger, Genotype influences in vivo dopamine transporter availability in human striatum, Neuropsychopharmacology 22 (2000) 133–139.

[76] J. Mill, P. Asherson, C. Browes, U. D'Souza, I. Craig, Expression of the dopamine transporter gene is regulated by the 3′ UTR VNTR: evidence from brain and lymphocytes using quantitative RT-PCR, Am. J. Med. Genet. 114 (2002) 975–979.

[77] S.H. VanNess, M.J. Owens, C.D. Kilts, The variable number of tandem repeats element in DAT1 regulates in vitro dopamine transporter density, BMC Genet. 6 (2005) 55.

[78] C.H. Van Dyck, R.T. Malison, L.K. Jacobsen, J.P. Seibyl, J.K. Staley, M. Laruelle, R.M. Baldwin, R.B. Innis, J. Gelernter, Increased dopamine transporter availability associated with the 9-repeat allele of the SLC6A3 gene, J. Nucl. Med. 46 (2005) 745–751.

[79] L.K. Jacobsen, J.K. Staley, S.S. Zoghbi, J.P. Seibyl, T.R. Kosten, R.B. Innis, J. Gelernter, Prediction of dopamine transporter binding availability by genotype: a preliminary report, Am. J. Psychiatry 157 (2000) 1700–1703.

[80] S.R. Jones, R.R. Gainetdinov, M. Jaber, B. Giros, R.M. Wightman, M.G. Caron, Profound neuronal plasticity in response to inactivation of the dopamine transporter, Proc. Natl. Acad. Sci. USA 95 (1998) 4029–4034.

[81] F. Karoum, S.J. Chrapusta, M.F. Egan, 3-Methoxytyramine is the major metabolite of released dopamine in the rat frontal cortex: reassessment of the effects of antipsychotics on the dynamics of dopamine release and metabolism in the frontal cortex, nucleus accumbens, and striatum by a simple two pool model, J. Neurochem. 63 (1994) 972–979.

[82] M. Matsumoto, C.S. Weickert, S. Beltaifa, B. Kolachana, J. Chen, T.M. Hyde, M.M. Herman, D.R. Weinberger, J.E. Kleinman, Catechol O-methyltransferase (COMT) mRNA expression in the dorsolateral prefrontal cortex of patients with schizophrenia, Neuropsychopharmacology 28 (2003) 1521–1530.

[83] M. Matsumoto, C.S. Weickert, M. Akil, B.K. Lipska, T.M. Hyde, M.M. Herman, J.E. Kleinman, D.R. Weinberger, Catechol O-methyltransferase mRNA expression in human and rat brain: evidence for a role in cortical neuronal function, Neuroscience 116 (2003) 127–137.

[84] L. Yavich, M.M. Forsberg, M. Karayiorgou, J.A. Gogos, P.T. Mannisto, Site-specific role of catechol-O-methyltransferase in dopamine overflow within prefrontal cortex and dorsal striatum, J. Neurosci. 27 (2007) 10196–10209.

[85] S.R. Sesack, V.A. Hawrylak, C. Matus, M.A. Guido, A.I. Levey, Dopamine axon varicosities in the prelimbic division of the rat prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter, J. Neurosci. 18 (1998) 2697–2708.

[86] B.H. Schott, C.I. Seidenbecher, D.B. Fenker, C.J. Lauer, N. Bunzeck, H.G. Bernstein, W. Tischmeyer, E.D. Gundelfinger, H.J. Heinze, E. Duzel, The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging, J. Neurosci. 26 (2006) 1407–1417.

[87] B. Knutson, G.W. Fong, S.M. Bennett, C.M. Adams, D. Hommer, A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI, Neuroimage 18 (2003) 263–272.

[88] A. Bertolino, G. Blasi, V. Latorre, V. Rubino, A. Rampino, L. Sinibaldi, G. Caforio, V. Petruzzella, A. Pizzuti, T. Scarabino, M. Nardini, D.R. Weinberger, B. Dallapiccola, Additive effects of genetic variation in dopamine regulating genes on working memory cortical activity in human brain, J. Neurosci. 26 (2006) 3918–3922.

[89] X. Caldu, P. Vendrell, D. Bartres-Faz, I. Clemente, N. Bargallo, M.A. Jurado, J.M. Serra-Grabulosa, C. Junque, Impact of the COMT Val108/158 Met and DAT genotypes on prefrontal function in healthy subjects, Neuroimage 37 (2007) 1437–1444.

[90] J-C. Dreher, P. Kohn, B. Kolachana, D.R. Weinberger, K.F. Berman, Variation in dopamine genes influences responsivity of the human reward system, Proc. Natl. Acad. Sci. USA (2009) Sci USA 106:15106–15111.

[91] J. Yacubian, T. Sommer, K. Schroeder, J. Glascher, R. Kalisch, B. Leuenberger, D.F. Braus, C. Buchel, Gene-gene interaction associated with neural reward sensitivity, Proc. Natl. Acad. Sci. USA 104 (2007) 8125–8130.

[92] B. Knutson, S.E. Gibbs, Linking nucleus accumbens dopamine and blood oxygenation, Psychopharmacology (Berl) 191 (2007) 813–822.

[93] J-C. Dreher, A. Meyer-Lindenberg, P. Kohn, K.F. Berman, Age-related changes in midbrain dopaminergic regulation of the human reward system, Proc. Natl. Acad. Sci. USA 105 (2008) 15106–15111.

[94] Y.C. Chen, W.R. Galpern, A.L. Brownell, R.T. Matthews, M. Bogdanov, O. Isacson, J.R. Keltner, M.F. Beal, B.R. Rosen, B.G. Jenkins, Detection of dopaminergic neurotransmitter activity using pharmacologic MRI: correlation with PET, microdialysis, and behavioral data, Magn. Reson. Med. 38 (1997) 389–398.

[95] G.K. Murray, P.R. Corlett, L. Clark, M. Pessiglione, A.D. Blackwell, G. Honey, P.B. Jones, E. T. Bullmore, T.W. Robbins, P.C. Fletcher, Substantia nigra/ventral tegmental reward prediction error disruption in psychosis, Mol. Psychiatry 13 (239) (2008) 267–276.

[96] W. Schultz, P. Dayan, P.R. Montague, A neural substrate of prediction and reward, Science 275 (1997) 1593–1599.

[97] R.S. Sutton, A.G. Barto, Reinforcement learning: an introduction, IEEE Trans Neural Netw 9 (1998) 1054.