Building Bridges between Neuroscience and Complex Decision Making

Woo-Young Ahn[1], Ryan K. Jessup[2], and Jerome R. Busemeyer[3,CA]

(1 Department of Psychological and Brain Sciences, Indiana University)

(2 College of Business Administration, Abilene Christian University)

(3 Department of Psychological and Brain Sciences, Indiana University)

Abstract: Animals and humans constantly evaluate and make choices to maximize reward and avoid punishment (value-based decision making). Recently dopamine has been known to play a major role in reward processing, and the cognitive and neural processes of value-based decision making have been widely studied using monkey electrophysiology and human functional resonance imaging (fMRI). We systematically review the previous findings involved in two critical stages of value-based decision making: valuation and action selection. We describe that human and monkey studies have produced consistent neural substrates for valuing appetitive and aversive events. We show sequential sampling models can be successfully applied to both behavioral and electrophysiological data of action selection, which suggests that they can provide a critical theoretical bridge between neuroscience and behavior.

CA email: jbusemey@ indiana. edu

Introduction

How do humans and animals make decisions and what cognitive and neural processes contribute to decision making? These are fundamental yet complex problems in cognitive neuroscience. The inherently interdisciplinary nature of the field has attracted people in various fields including psychology, economics, computer science, and neuroscience. Scientists have begun to examine the neural basis of decision making, specifically value-based decision-making, that is, the influence of appetitive (e.g., reward) and aversive (e.g., punishment) events[①] on choice.

Figure 1 shows a simple framework illustrating stages for making value-based decisions (Rangel, Camerer, & Montague, 2008). Although most decision stages are not quite so explicitly separate, it will be helpful to briefly step through these while working through an example involving a teenager who is offered a cigarette to smoke.

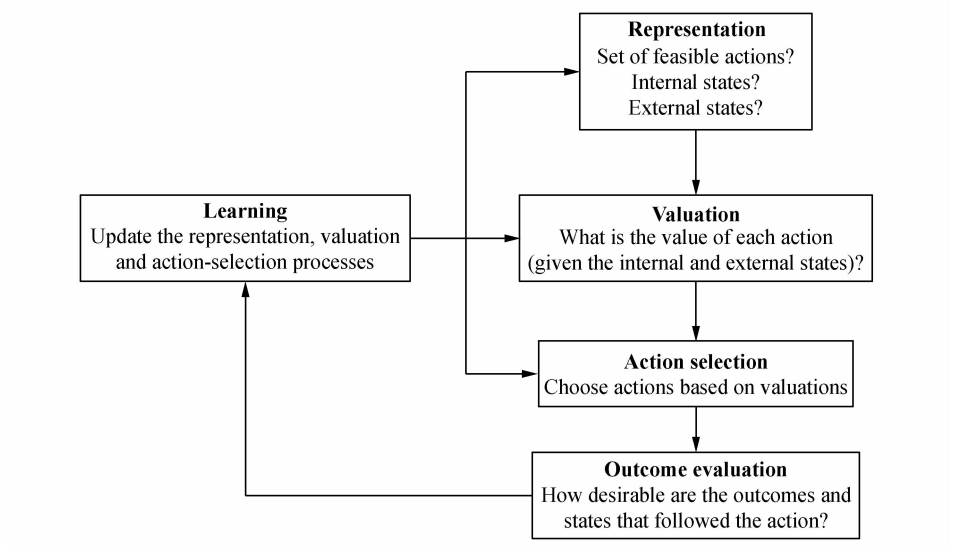

Figure 1 Conceptual framework illustrating processes involved in value-based decision making. Adapted with permission from Rangel, Camerer, & Montague (2008).

① A reward is defined as any object or event which induces approach behavior, subjective feelings of pleasure and hedonia, or increases the frequency and intensity of such behavior (Schultz, 2007). A punishment has the opposite meaning: any object or event which induces withdrawing behavior, negative feelings, or decreases the frequency and intensity of such behavior.

The representation stage signifies available actions and the current context. In our example, the available actions are accepting or declining the offer, and the context might be at a late night party surrounded by friends or, alternatively, in the presence of the teenager's parents in the middle of the day. Next, the available actions can be affectively evaluated within their context. For example, accepting the offer to smoke may be more positively valued at a late night party in front of people the teenager wants to impress but may not be tempting at all in front of the teenager's own parents. Based on these affective evaluations an action is selected. In general, the action with the most positive affective evaluation is thought to be selected, however, noise inherent in neural activity as well as unknown features (i.e., from an observer's perspective) generally render action selection probabilistic, such that the highest value is most likely—but not certain—to be selected. After an action is selected and the outcome experienced, the result is evaluated. For example, if the teenager accepts the offer to smoke while at the party and embarrassingly chokes on the smoke, he or she will probably evaluate the experience as much more negative than before. This evaluation is then used to update aspects of the previous three stages. In our example, if presented with a follow-up offer to smoke at a subsequent late night party, the teenager may now have a much more negative evaluation of accepting the offer to smoke. In what follows, we will focus most prominently on the valuation and action selection stages of the value-based decision making process.

Several papers have reviewed related topics (Berridge & Aldridge, 2008; Cohen, 2008; Daw & Doya, 2006; Dayan & Daw, 2008; Niv, 2009; O'Doherty & Bossaerts, 2008; Platt & Huettel, 2008; Rangel, et al., 2008; Schultz, 2007) and this chapter aims to extend previous reviews with more recent studies and summarize the findings in a different and systematic way. First, we will provide a brief introduction to the dopamine system, which plays a major role in reward processing. Second, we will examine the neural substrates of value-based decision making. This will be separated into two major sections: monkey electrophysiology and human functional magnetic resonance imaging (fMRI). These two sections will consider decision making involving appetitive and aversive events separately. Third, we will briefly summarize previous research on sequential decision-making and its neural substrates. Fourth, we will briefly review a relevant model which was initially created for behavioral choice but is applicable to neural data as well.

Dopamine System

Although dopamine (DA) is involved in multiple functions such as movement, cognition, motivation and learning (Schultz, 2007; Wise, 2004), its role in reward processing has been recently highlighted. Anatomically, DA-containing nerve cells are localized in the midbrain, the diencephalon (e. g., thalamus and hypothalamus), and the olfactory bulb. There are three main DA systems—the nigrostriatal system, the mesolimbic system, and the mesocortical system. The nigrostriatal system originates from cells in the substantia nigra (SN) and projects to the caudate nucleus and the putamen. It is mostly closely associated with motor function. The mesolimbic system originates in the ventral tagmental area (VTA) and projects to the nucleus accumbens and the olfactory tubercle, but also to the septum, amygdala, and hippocampus. The mesocortical system also originates in the VTA but projects to the prefrontal cortex (PFC), cingulate cortex and perirhinal cortex (Arias-Carrion & Popel, 2007). Because of substantial overlap between the mesolimbic and the mesocortical systems, they are often referred to as the mesocorticolimbic dopamine system (Wise, 2004). Also, DA neurons receive signals from diverse other areas including the medial prefrontal cortex, the nucleus accumbens, amygdala (the central nucleus), the lateral hypothalamus, the lateral habenula, the ventral pallidum, and the noradrenergic locus coerulues and they interact as a dynamic regulated network (Durstewitz, Kelc, & Gunturkun, 1999).

Neurotransmitters act via receptors and at least five distinct receptor subtypes exist for DA—the D1-like receptor subtype (D1 and D5) and the D2-like receptor subtype (D2, D3, and D4). The densities of D1 and D2 receptors are highest in the basal ganglia (BG; consisting of the caudate nucleus, putamen, globus pallidus pars medialis), substantia nigra and olfactory tubercle in human brain (Camps, Cortes, Gueye, Probst, & Palacios, 1989; Cortes, Camps, Gueye, Probst, & Palacios, 1989). In the PFC, the densities of D1 receptors are 4 to 7 times higher than those of the D2 receptor(Hall et al., 1994). In contrast, Camps et al. (1989) found that the densities of the D2 receptors were predominantly higher in BG compared to PFC. Interestingly, the densities of D1 (Wang et al., 1998) and D2 (Seeman et al., 1987; Wang et al., 1995) receptors in human brain were found to decrease with age, which may shed insights to age-related changes in the reward system (Dreher,

Meyer-Lindenberg, Kohn, & Berman, 2008).

The electrophysiological properties of DA cells alternate between a state of phasic (transient) release and tonic (sustained) release. The phasic DA release is associated with event-specific information, and is caused by dopamine neuron firing and rapidly terminated by reuptake transport into the DA terminals before it diffuses extrasynaptically. Meanwhile, the tonic DA release is regulated by prefrontal cortical afferents and may control the background level of DA in the synapse. Some evidence suggests that tonic DA release regulate the phasic DA response(Grace, 1991).

The tonic release is considered to be primarily modulated by D1 receptors (Durstewitz, et al., 1999; Durstewitz & Seamans, 2002) while the phasic release may be modulated by D2 receptors. However, because the level of tonic DA release regulates the response of phasic release in the opposite direction, the effects mediated by D1 or D2 can indirectly influence the effects mediated by the other receptor type (Cohen, Braver, & Brown, 2002). Cohen et al. (2002) proposed an integrative tonic/phasic model of DA, in which tonic D1 release is associated with maintaining working memory (WM) and phasic D2 release is associated with updating and learning new information. The model explains previous findings that low DA level (i.e., low tonic D1, thus low WM maintenance) produces impulsive behavior while high DA level (i.e., high tonic D1, thus low phasic D2) results in perseverative behavior. Frank and colleagues proposed a more sophisticated neurocomputational framework of DA and BG with specific roles of D1 and D2 receptors (Frank & Claus, 2006; Frank & O'Reilly, 2006; Frank, Woroch, & Curran, 2005). Specifically, it is claimed that activity in the direct projection pathway from the BG facilitates the most adaptive cortical response (go pathway), while activity in the indirect projection pathway inhibits the inappropriate cortical response (no-go pathway). Furthermore, D1 and D2 receptors are known to drive go and no-go pathways, respectively (for more details, see Frank & O'Reilly, 2006).

Neural Substrates of Valuation and Outcome-evaluation

Then the next question is what brain regions are involved in value-based decision-making, especially in the stages of valuation and outcome evaluation. We will summarize previous findings of monkey electrophysiology and human studies separately.

Monkey Electrophysiology Studies

Appetitive event signals

Most DA neurons (60%—80%) in SN and VTA show phasic responses to unexpected primary rewards such as food and liquid as well as to reward-predicting cues (Schultz, 2007). A series of studies done in Schultz's lab in early 1990s demonstrated that DA responses are not equal to the motivational value of primary rewards (Ljungberg, Apicella, & Schultz, 1992; Romo & Schultz, 1990; Schultz, Apicella,

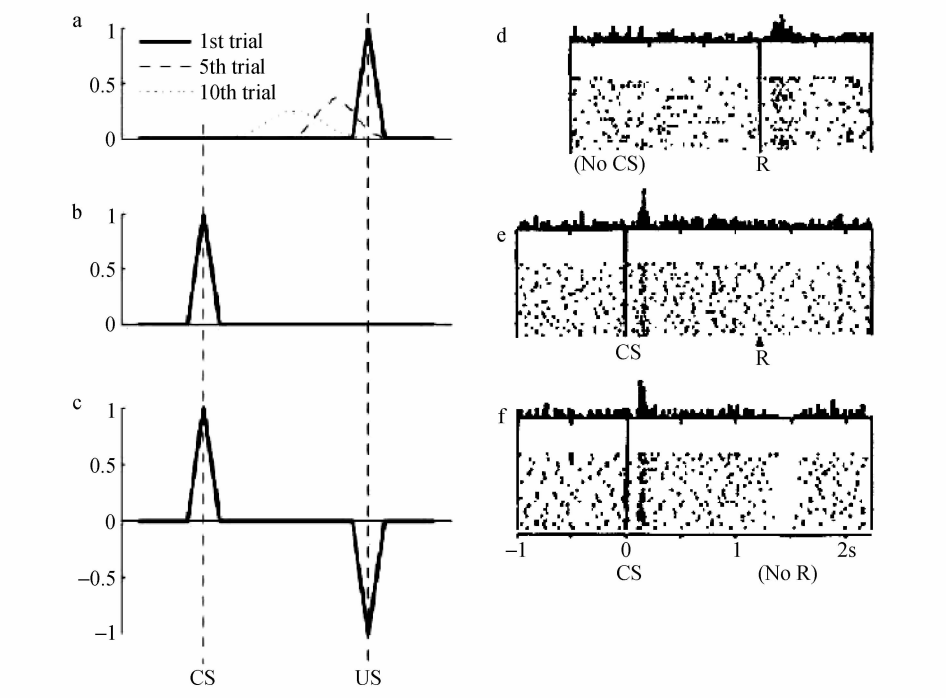

Figure 2 Phasic firing patterns of DA neurons in a Pavlovian conditioning task. (a-c) Predictions from the temporal-difference learning algorithm. (d-f) actual DA responses adapted from Schultz et al. (1993). (a & d) Before learning, DA neurons respond to a reward (R) or unconditioned stimulus (US) itself (appetitive fruit juice) but not to a reward-predicting cue (conditioned stimulus, CS). (b & e) After learning, DA neurons respond to CS, but not to a delivered reward because the reward is already fully predicted by CS and not surprising (c & f). When CS predicts a reward but none occurs, the activation of DA neurons is depressed at the expected time of reward delivery. The temporal difference learning algorithm well describes these DA responses. Adapted with permission from Niv (2009).

& Ljungberg, 1993). During the initial learning stage in such studies, DA neurons in the VTA and the SN responded to primary rewards, but not to cues that predicted rewards, which was consistent with the "dopamine as reward" hypothesis (see Figure 2). After conditioning occurred, however, DA neurons responded to reward-predicting cues, but not to fully-predicted rewards. Also, even after conditioning DA neurons responded to unexpected rewards, and unexpected reward omission induced depression in activity. Subsequently, It was demonstrated that such transfer of the DA responses can be well described by a reinforcement learning model, called the temporal-difference reinforcement learning algorithm (Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997) or a neural network model (Brown, Bullock, & Grossberg, 1999).

The results suggest that DA neurons encode the degree of surprise or prediction errors, the difference between the observed reward and its predicted reward value, called the prediction error (PE). Furthermore, the activity of DA neurons appears to carry more detailed reward information such as the expected value. For instance, it has been shown that when visual cues predicted different juice reward magnitudes and probabilities, the responses of DA neurons increased with increasing expected values or probabilities of rewards (Fiorillo, Tobler, & Schultz, 2003; Morris, Arkadir, Nevet, Vaadia, & Bergman, 2004; Tobler, Fiorillo, & Schultz, 2005). The neuronal activities in dopamine-projected regions such as thePFC(Leon & Shadlen, 1999; Padoa-Schioppa & Assad, 2006; Rosenkilde, Bauer, & Fuster, 1981; Thorpe, Rolls, & Maddison, 1983), the amygdala(Belova, Paton, Morrison, & Salzman, 2007; Paton, Belova, Morrison, & Salzman, 2006), and the caudate nucleus (Hikosaka, Sakamoto, & Usui, 1989a, 1989b, 1989c) are also modulated with expected reward signals.

Aversive event signals

Although it is widely accepted that DA neurons encode appetitive outcomes and positive prediction errors, it has remained unclear whether they encode aversive outcomes (e. g., airpuffs) and whether negative prediction errors originate from DA neurons. While some studies (Mirenowicz & Schultz, 1996; Ungless, Magill, & Bolam, 2004) reported that aversive outcomes and their predicting cues homogeneously inhibit DA activity, other studies (e. g., Chiodo, Antelman, Caggiula, & Lineberry, 1980; Guarraci & Kapp, 1999; Schultz & Romo, 1987) reported that some DA neurons were excited and some were inhibited upon aversive events. Alternatively,

Daw and colleagues have proposed that background firing rate of DA neurons might reflect the overall rate of aversive events and serotonin may be responsible for aversive signals (Daw, Kakade, & Dayan, 2002).

Findings from more recent studies suggest two new possibilities. First, there might be two types of DA neurons; A recent study found that some DA neurons were uniformly excited by appetitive events (juice) and inhibited by aversive events (air-puffs) while a separate set of DA neurons were phasically excited by both appetitive and aversive events (Matsumoto & Hikosaka, 2009). In addition, the tonic level (i.e., background firing rate) was not significantly different between the appetitive blocks and the aversive blocks in the study, which is inconsistent with the proposal of Daw et al. (2002). Second, the lateral habenula, which is a part of the epithalamus, might be a major source of negative reward signals in DA neurons. Matsumoto & Hikosaka (2007) showed that DA neurons were excited by positive reward signals and inhibited by negative reward signals, replicating previous finding. Meanwhile, habenula neurons showed the opposite pattern and their activity started earlier than that of DA neurons, suggesting inhibitory inputs from habenula neurons contribute to DA neuron activities.

In summary, accumulating evidence suggests that DA plays a critical role in processing appetitive events. DA responses encode not just appetitive events but even more complex information of rewards such as the expected value of rewards and prediction error signals. The role of DA neurons in aversive events remain unclear, but recent studies suggest there might be more than one type of DA neuron; moreover, habenula neurons appear to interact with DA neurons when processing aversive events.

Human fMRI Studies

Recent development of blood-oxygen-level dependent (BOLD) fMRI has enabled researchers to explore brain areas associated with specific decision making processes in humans. Although the BOLD fMRI provides indirect measures of neural activity and its signals cannot be interpreted as DA or neural activity itself (Knutson & Gibbs, 2007), it is one of the best tools to study the neural substrates of decision making processes.

Appetitive event signals

Overall, findings from previous human fMRI studies on the neural correlates of

reward signals are in line with those from monkey electrophysiology studies. DA terminal areas including the striatum, the orbitofrontal cortex (OFC), and the amygdala have been associated with the encoding of reward signals. For instance, Gottfried et al. (2003) found reward-related activity in the OFC②, the ventral striatum, and the amygdala using a simple Pavlovian task. Also, abstract rewards such as money (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000), beautiful faces (Aharon et al., 2001), cultural objects (Erk, Spitzer, Wunderlich, Galley, & Walter, 2002), humor (Mobbs, Greicius, Abdel-Azim, Menon, & Reiss, 2003), and romantic love (Aron et al., 2005) modulated the activity in ventral striatum, the OFC or other DA-projection regions. To examine the neural correlates of expected and experienced outcomes, the monetary incentive delay (MID) task has been widely used(Knutson, Adams, Fong, & Hommer, 2001; Knutson, Fong, Adams, Varner, & Hommer, 2001). In the MID task, participants first see cues that predict different monetary gains and losses and try to respond to a target (presented only for 160—260 msec) after a variable delay to earn or avoid the outcome. Task difficulty was set such that successful response rate would be approximately 66%. Studies using the MID task have shown that the activity of the ventral striatum increased with the magnitude of anticipated gain. Tom et al. (2007) used a gambling task and investigated the neural correlates of anticipated gains and losses. Participants played a gambling task in which they decided whether to accept or reject wagers that offered a 50/50 chance of gaining or losing money. It was found that the activity of the striatum—especially ventral striatum—and the mOFC increased with the magnitude of potential gains and decreased with the magnitude of potential losses. In another study using a guessing card task (Preuschoff, Bossaerts, & Quartz, 2006), the probability of winning on a trial was systematically modulated from 0 to 1 and expected reward was significantly correlated with activation in the ventral striatum, the dorsal striatum, the mPFC, anterior cingulate cortex (ACC), and the midbrain. All of these results suggest that the striatum (especially the ventral striatum) and the OFC encode the expected reward signals in BOLD fMRI.

② Some people differentiate the OFC and the medial prefrontal cortex (mPFC) while some do not. Wallis (2007) referred to the OFC as Brodmann area (BA) 10, 11, 13, 14 and 47/12 and the mPFC as BA 32 and 24. Roughly speaking, O'Doherty (2007) referred to the former as central areas of OFC and the latter as the medial OFC. In this review, we will use the notation of O'Doherty (2007) for simplicity.

The studies reviewed above used description-based tasks. Description-based decisions refer to a situation when individuals are presented with an explicit description of alternatives. Experience-based decisions refer to when individuals learn about alternatives through experience. The expected reward is learned in the latter case while it is merely given in the former case. An appropriate question, then, is what are the neural substrates used when learning about rewarding events? While the PE signals in monkey studies were found in the VTA and the SN, fMRI studies in humans have onlyfound such signals in the striatum and the OFC (e.g., McClure, Berns, & Montague, 2003; O'Doherty, 2003), presumably because of technical difficulties involved in imaging the brainstem. One would expect to find DA signals in the striatum because DA neurons project to the striatum and the BOLD response is believed to reflect the presynaptic input more than spiking output (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). However, a recent study (D'Ardenne, McClure, Nystrom, & Cohen, 2008) showed that BOLD responses can be measured in the VTA using new imaging techniques, thus supporting the existence of PE signals in the human VTA. Further evidence came from a combined pharmacological and fMRI study (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). While subjects treated with the dopamine agonist L-DOPA showed improved behavioral performance and increased BOLD signals for appetitive PEs, subjects treated with the dopamine antagonist haloperidol performed worse than L-DOPA treated subjects. Furthermore, appetitive PE signals in the ventral striatum were significantly higher in L-DOPA treated subjects compared to haloperidol treated subjects.

While the ventral striatum has been implicated for encoding PE signals in various tasks, the dorsal striatum is known to be important for learning action-outcome contingency. The dorsal striatum plays an important role in trial-and-error learning and habit formation(Yin & Knowlton, 2006) and research in Parkinson's disease patients further supports the link (Delgado, 2007; Schonberg et al., 2010). O'Doherty and colleagues showed that while the ventral striatum correlated with PE signals during both Pavlovian and instrumental conditioning, the dorsal striatum correlated with PE signals only in an instrumental task(O'Doherty et al., 2004). The results suggest that the dorsal striatum may behave like the "actor" of the actor-critic model (Sutton & Barto, 1998) and maintains action-outcome contingencies (i.e., the perception that their action determined the outcome) to guide actions. Another study from the O'Doherty group examined whether PE signals in the striatum can distinguish learn-

ers and nonlearners in a four-armed bandit task (Schonberg, Daw, Joel, & O'Doherty, 2007). PE signals in the dorsal striatum was significantly higher in learners compared to nonlearners and the magnitude of PE signals in the region correlated with behavioral performance index.

With respect to brain areas encoding expected future reward in more complex tasks such as an n-armed bandit task which requires goal-directed behavior (e.g., maximizing earnings), the OFC has been implicated in encoding the value of chosen actions. To find the neural correlates of value signals in such tasks, a method called "model-based fMRI analysis" is increasingly being used. In model-based fMRI, a mathematical model is fit to the behavioral data to infer the hidden states of cognitive processes and predictions derived from the model are correlated against fMRI data to determine brain areas related to the specific processes. This method has become popular in decision neuroscience because it provides insights into the neural correlates of predicted cognitive processes and can be useful to discriminate competing theories of brain function(O'Doherty, Hampton, & Kim, 2007). Generally, the OFC is known to be involved in goal-directed learning, encode the affective value of stimuli, and update such expectations (O'Doherty, 2007; Wallis, 2007). The OFC is highly interconnected with the amygdala and the ventral striatum (Carmichael & Price, 1995) as well as other parts of the PFC (Carmichael & Price, 1996). Studies using model-based fMRI analysis to estimate expected reward also revealed expectation-related activations in the mOFC (Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; Hampton, Adolphs, Tyszka, & O'Doherty, 2007; Hampton, Bossaerts, & O'Doherty, 2006; Kim, Shimojo, & O'Doherty, 2006; Tanaka et al., 2004).

Whether expected rewards and experienced (or evaluated) rewards are processed in the same or distinct neural systems has been an important topic in both animals and humans. The accumulated evidence in human studies suggests two major regions evaluate appetitive outcomes: the OFC and the dorsal striatum. Several BOLD fMRI studies have shown that the activity in the mOFC correlates with subjective experience of primary rewards in olfactory, auditory, gustatory and visual domains (O'Doherty, 2007; Rangel, et al., 2008). In addition, the evidence suggests that the outcome evaluation process of the mOFC is modulated by top-down cognitive processes (de Araujo, Rolls, Velazco, Margot, & Cayeux, 2005; McClure et al., 2004; Plassmann, O'Doherty, Shiv, & Rangel, 2008). The receipt of monetary payoffs is also implicated in the mOFC (Breiter, Aharon, Kahneman, Dale, & Shizgal,

2001; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Knutson, Fong, et al., 2001; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). The involvement of the human dorsal striatum in outcome evaluation was first showed by Delgado and colleagues (Delgado, Locke, Stenger, & Fiez, 2003; Delgado, et al., 2000). They used a card-guessing game in which participants guess if the number on a card would be greater or less than the number 5. The number could be any integer from 1 to 9, so there was a 50/50 chance of winning on each independent trial. At the time of feedback, BOLD signals in the dorsal and the ventral striatum increased on rewards (monetary gains) and decreased upon punishments (monetary losses). Only in the dorsal striatum, however, the magnitude of change was modulated by the amount of monetary gains/losses. A subsequent study investigated why the dorsal striatum was implicated in outcome evaluation(Tricomi, Delgado, & Fiez, 2004). Using three elegantly designed experiments, the effects of reward/punishment anticipation and action-outcome contingency on the activation of the dorsal striatum (the caudate nucleus) were examined. They showed that the dorsal striatum activation was observed only when action-outcome contingency existed. The results suggest that when participants' responses are irrelevant with the outcomes, the dorsal striatum was not activated (Berns, McClure, Pagnoni, & Montague, 2001; Breiter, et al., 2001; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003).

In summary, this section reviewed findings from human fMRI studies on expected and experienced appetitive events. The results suggest that the ventral striatum and the OFC have been repeatedly implicated for encoding expected reward signals. Consistent with monkey studies, the striatum appears to encode PE signals and the activity in the dorsal striatum might be associated with action-outcome contingency. With respect to neural correlates of experienced appetitive events, previous literature suggests that the OFC and the dorsal striatum might encode the events. However, it is still unclear whether the striatum, the OFC, or both process the affective properties of outcomes when action-outcome contingency exists. For instance, O'Doherty et al. (2001) used a similar guessing task and Daw et al. (2006) used a four-armed bandit task that also engender the action-outcome contingency, and found that the activity of the mOFC, but not the dorsal striatum, correlated with the magnitude of the obtained outcomes. Meanwhile, some studies (Filoteo et al., 2005) including those by Delgado and colleagues failed to find activation in the mOFC whereas the dorsal striatum correlated with feedback outcomes. Possibly differing amounts of orbital signal loss due to

differing scanners and analysis methods engender the mixed results for OFC activity.

Aversive event signals

Next, we turn to researcher on aversive events. Several human studies examined the neural correlates of physical punishment such as pain and found that the activity of insula, ACC, and the thalamus correlated with physical intensity (Peyron et al., 1999; Peyron, Laurent, & Garcia-Larrea, 2000; Wager et al., 2004). With respect to neural correlates of secondary punishments such as monetary losses, findings from fMRI literature are unclear and mixed like the findings from monkey studies reviewed previously. ③ The central issue in these studies is whether a single, unitary system encodes both appetitive and aversive events or separate systems encode each.

With respect to negative PE signals for punishments or absence of rewards, the fMRI results are also mixed. For example, some studies reported that the ventral striatum encodes both positive and negative PE signals (McClure, et al., 2003; O'Doherty, et al., 2003), Seymour et al. (2007) reported that the striatum encodes both positive and negative PE signals but they are functionally segregated in subregions (more anterior regions for encoding positive PEs and more posterior regions for encoding negative PEs). Delgado et al. (2000) reported that the dorsal striatum evaluated both the appetitive and aversive outcomes. Similarly, some studies (Breiter et al., 2001; O'Doherty et al., 2001) reported that the OFC encodes both appetitive and aversive outcomes. Tom et al. (2007) showed that BOLD signals in the nucleus accumbens and the PFC increased when anticipating potential gains and decreased when anticipating potential losses. They found no activation upon anticipated potential losses in amygdala or anterior insula, which have been associated with negative emotion and losses (Breiter et al., 2001; Kahn et al., 2002; Kuhnen & Knutson, 2005). The authors argued that it is due to the difference between "anticipated utility" and "decision utility". However, often those two utilities are closely related and this hypothesis needs further clarification (Berridge & Aldridge, 2008).

All these studies suggest that a single system encodes both appetitive and aversive values. Meanwhile other studies support the claim that multiple brain areas separately encode appetitive and aversive events. For instance, Yacubian et al. (2006) showed

---

③ One thing to keep in mind is humans have a unique brain structure (anterior insula) for experiencing subjective feelings of pain, therefore rat or monkey studies cannot inform this specific system and vice versa (Craig, 2009).

that while the ventral striatum encodes appetitive events, the amygdala encodes the aversive events. Knutson & Gibbs (2007) also observed that the insula encodes aversive events. Importantly, a recent study used a novel method to image the habenula in humans using fMRI (Salas, Baldwin, de Biasi, & Montague, 2010) and the fMRI findings agree with those from animal electrophysiology (Matsumoto & Hikosaka, 2007).

In summary, fMRI findings have produced mixed results regarding whether there is a single or multiple brain regions for encoding aversive event signals and it is partly due to the indirect nature of BOLD imaging. However, the fMRI results of Salas et al. (2010) confirm the role of habenula in encoding aversive event signals.

Action Selection—Decision Field Theory

Neuroscientists are just beginning to understand how values are actually translated into actions. An emerging literature using electrophysiological recording techniques in monkeys in visual and attention research suggests that expectations for actions evolve over time by accumulating information from the environment about possible actions. These research methods have been extended to the neural basis of decision-making (for reviews, see Gold & Shadlen, 2001, 2002; Platt, 2002; Schall, 2003). The sluggish BOLD response used as a dependent variable in fMRI research is unable to detect the time-course of this evolution of a decision because the timescale of a decision is very short. The neurophysiological studies show that both the choice and decision time of the monkeys were accurately predicted using what are known as sequential sampling models (Figure 3). Sequential sampling models used in behavioral choice research can be mathematically fit to behavioral choice and response time data, and then used to generate predictions for neural activation patterns obtained from electrophysiological recordings of the motor regions such as frontal eye fields (FEF) during choice (Ratcliff & Smith, 2004). Although sequential sampling models are mathematical models, rather than fine grain neural models, they provide a critical theoretical bridge between neuroscience and behavior (Wang, 2008).

A typical task in this electrophysiological recording paradigm involves showing monkeys a display of dots moving in a mostly random fashion, with only a slight trend in one direction (e.g., up or down). The amount of trend or coherence in the dot motion can be manipulated from strong, making the decision easy, to non-existent, reducing the decision to chance. Monkeys are trained to detect the dot motion and

saccade from a central fixation point to an upper location when they perceive that the dots are moving mostly upwards, and to a lower location when they believe the dots are moving mostlydownwards.

Several prominent findings from these tasks are relevant here. First, while recording from middle temporal (MT) area—an area known to detect the motion of objects-Ditterich et al. (2003) were able to simultaneously stimulate "up" cells in MT. They observed that this microstimulation of up cells (added evidence) resulted in faster decisions when the dots were moving upwards, consistent with the predictions form a sequential sampling model during an easy or highly discriminable choice. When the dots were moving downwards while simultaneously stimulating up cells (inconsistent evidence), choices were slower as predicted by sequential sampling models during difficult choices. Such an effect is consistent with the idea that MT is detecting evidence in a "bit by bit" manner. By microstimulating the "up" cells in area MT, the researchers were essentially tricking the brain so that it perceived additional evidence that the dots were moving upwards

Second, in a separate study using a similar task and procedure, FEF neurons—an area which controls saccadic eye movements—were microstimulated during simultaneous recording(Gold & Shadlen, 2000). Here, the stimulation forced a rightward saccade. However, the rightward saccade deviated systematically such that as viewing duration or the motion coherence of the dots increased, the more the saccade deviated in the direction of the dot motion. This finding is consistent with the notion that the visual information which is used to evaluate an action is being accumulated in the FEF neurons which will command the action. In this case, the more that information favoring an action has been accumulated the further the forced saccade deviated in the direction of the favored action.

Third, looking at the activity in the lateral intraparietal (LIP) area,an attention region which has extensive outputs to the FEF and superior colliculus—two areas which are crucial for saccadic eye movements—provides evidence that information is accumulated to a threshold before an action is selected(Gold & Shadlen, 2007). When LIP activity is aligned on the onset of a trial, it reveals that for all levels of coherence, increased viewing duration results in increased LIP activity for the cells representing the eventually chosen action. Moreover, at each individual time point but across coherence levels, the greater the amount of motion coherence in the dots—and hence, the easier the decision—the greater the LIP activity for the cells representing the eventually chosen action. This result is also consistent with the notion of accumu-

lation of information as described above. However, when the LIP activity is aligned at the onset of the saccade, it did not differ between viewing duration or motion coherence, indicating that accumulated information in the LIP integrates to relatively fixed threshold, at which point the monkey has enough information to select an action. Interestingly, these results are not limited to the visual domain but have also been shown in somatosensory and manual cortical areas using somatosensory tasks with manual actions (Romo et al., 1999, 2002), as well as tasks in which the reward probabilities and magnitudes differ between actions (Platt & Glimcher, 1999). A recent study in Newsome's group showed that sensory and reward information are combined in LIP (Rorie, Gao, McClelland, & Newsome, 2010). Interestingly, the effects of payoffs influenced the starting point of the sequential sampling model, but not the drift rate of the model. Taken as a whole, these results suggest that the findings from this decision making paradigm are rather broad and that the accumulatory areas are perhaps accumulating value as opposed to merely information.

Decision Field Theory (DFT, Figure 3) uses this same type of sequential sampling process to model more complex decisions such as choosing among consumer products, or choosing between uncertain and risky courses of action (Busemeyer, Jessup, Johnson, & Townsend, 2006; Busemeyer & Townsend, 1993; Johnson & Busemeyer, 2005; Roe, Busemeyer, & Townsend, 2001). According to DFT, the decision maker deliberates over each course of action by thinking about possible outcomes and feeling the anticipated consequences of alternative actions. These comparisons are accumulated over time to form a preference state (Figure 3B); when the trajectory for an action reaches a threshold, that action is selected, providing predictions for both choice and decision time. DFT can also generate both neurophysiological predictions (Figure 3B, 3C) and model-based time-series regressors to predict fMRI data during decision making with humans. Thus DFT provides a critical theoretical bridge between neural models from neuroscience and complex decision making behavior of interest to decision scientists.

The *threshold bound* for the decision process, represented by the upper horizontal line (Figure 3B), is the key parameter for controlling speed and accuracy tradeoffs. Impulsive individuals may tend to use lower thresholds, while perspicacious individuals may tend to use higher thresholds. Although DFT is a dynamic and stochastic model, this dynamic process accomplishes the same objectives as a subjective expected utility (SEU) model (Tversky & Kahneman, 1992). By integrating the sequentially

sampled evaluations of each action across time, the preference state evolves into an estimate of the SEU of each action, plus noise from sampling that gets averaged out with time. Furthermore, DFT uses the same parameters as used by expected utility models. Affective values and attention weights in DFT are equivalent to utilities and decision weights, respectively, in SEU. The threshold parameter in DFT is similar to the choice probability parameter used in logistic models of choice in traditional decision theories. All parameters of DFT can be estimated from the choice probabilities and these same

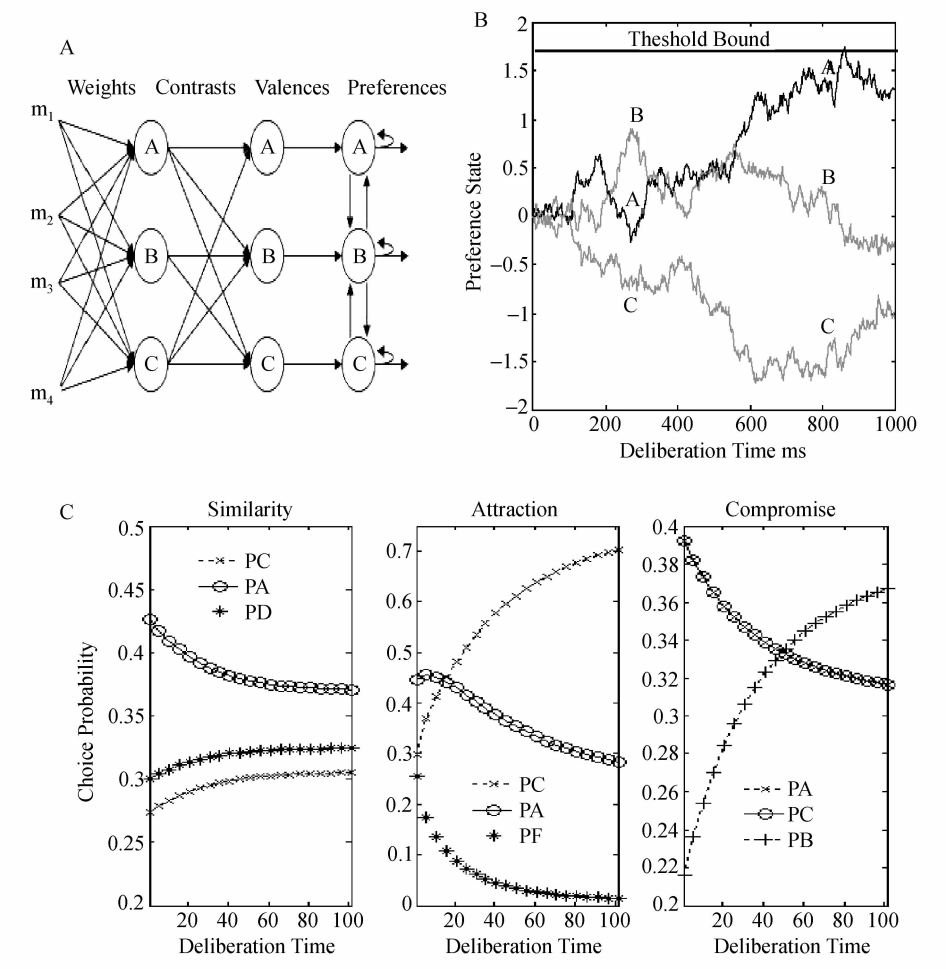

Figure 3 Decision Field Theory. (A) Network model of DFT. (B) DFT decisions occur when simulated neural activity reaches a threshold boundary. (C) DFT predicts specific, sometimes non-monotonic activity trajectories and rates of increase in monkey neurons.

parameters can be used to predict decision time and neural activation across time. Thus DFT is more testable than traditional decision theories because it predicts decision time and neural activation across time without estimating any new parameters.

Advantages of DFT in Context Effects

A great challenge for economic utility theories is that preferences are highly dependent on the context provided by the set of options presented for choice. For example, a teenager may change preferences between cigarette smoking or not if marijuana is added as a third option, making cigarettesthe lesser of two evils. These contextually induced preference reversals involve violations of a rational principle called *independence from irrelevantalternatives*. According to this principle, if option A is chosen more frequently than option C in a choice set that includes only $[A, C]$, then A should be chosen more frequently over C in a larger choice set $[A, B, X]$, that includes a new option X. However, empirical evidence points to at least three direct violations of this principle (Figure 3C). The first violation is produced by what is called the similarity effect(Tversky, 1972; Tversky & Sattath, 1979). In this case, suppose C is chosen more frequently than A in a binary choice. But then a new option D is added to the choice set, which is designed to be similar to and competitive with option C. This new option D steals probability from C but leaves the probability for A intact so that A becomes more popular than C in the triadic choice. The similarity effect rules out all simple scalable utility models (e.g., the softmax rule), but it can be explained by a heuristic choice model called the elimination by aspects (EBA) model (Tversky, 1972).

The second violation is produced by what is called the attraction effect (Huber, Payne, & Puto, 1982; Huber & Puto, 1983; Simonson, 1989) In this case, again suppose A is chosen more frequently than C in a binary choice. But then a new option F is added to the choice set, which is designed to be similar to but defective compared to C. This new option F enhances the probability of choosing C above its previous binary choice probability. In this second case, the new option helps rather than hurts the similar option. The attraction effect is important because it violates another rational principle called *regularity*, which states that adding an option to the set can never increase the popularity of one of the original options from the subset. The EBA model satisfies regularity, and therefore it cannot explain the attraction effect(Tversky, 1972). The third violation is produced by what is called the compromise effect

(Simonson, 1989; Tversky & Simonson, 1993). In this case, suppose option *A* is chosen over *B* more frequently in a binary choice. But then a new option *C* is added to the choice set, which makes option *B* appear as a compromise between two extremely different options *A* and *C*. Now the compromise option *B* becomes most popular. Thus adding an extreme option *C*—which turns option *B* into a compromise—reverses the preference orders obtained between binary and triadic choice. The compromise effect is interesting because it rules out another heuristic choice rule called the lexicographic (LEX) or "take the best" strategy. According to the LEX strategy, individuals should never choose the compromise option.

Can a single theory account for similarity, attraction, and compromise effects, using a common set of assumptions and a single set of parameters? Traditional utility models, including simple scalable and random utility models fail, and so do heuristic models such as the lexicographic or the elimination by aspects model (Roe, et al., 2001). However, DFT provides a robust and comprehensive account for all three effects (Figure 3C). Most importantly, except for a single lateral inhibitory parameter, all of the parameters are fixed by the binary choices, allowing new testable predictions for all of the triadic choices. For example, DFT must predict that the attraction and compromise effects grow stronger with longer deliberation times. If this prediction fails, the theory fails. This is bold because it predicts that longer deliberation produces what economists think is more irrational behavior. Nevertheless, this bold prediction was confirmed in experiments by Dhar et al. (2000).

Conclusion

Over the past decade, our knowledge of the neural substrates of value-based decision making has greatly advanced. Accumulating evidence suggests that phasic DA responses play a critical role in encoding rewarding events. Human and monkey research on DA and its role in decision making has revealed several brain systems responsible for each specific process. This chapter aimed to summarize existing literature concerning the neural substrates for appetitive and aversiveevents. Specifically, we summarized findings separately for monkey and human fMRI studies, appetitive and aversive events, and expected and experienced values.

To briefly summarize the findings, human and monkey studies produced consistent results in that in both areas, appetitive events correlated with neural signals in the

brainstem or DA projecting areas. Accumulating evidence suggests that DA neurons encode appetitive events and the signals are projected to the striatum and the OFC; however, the role of DA in encoding aversive events still remain unclear and needs further research and replication of recent findings. Evidence from human fMRI studies suggests that the ventral striatum is involved in expected appetitive events in most paradigms while the OFC and the dorsal striatum correlated with experienced appetitive events. Also, the OFC activity appears to be more selective for complex and experience-based decisions, which is consistent with its role of integrating multiple sources of information.

Neural activation patterns from FEF, MT, and LIP suggest that information is accumulated over time to a threshold before an action is selected. Sequential sampling models have been successfully usedto theoretically model behavioral and electrophysiological data of action selection. DFT—a specific sequential sampling model—can generate predictions for both sets of data and furthermore, it can provide unified explanation of context effects given a common set of assumptions and a single set of parameters.

Still, there remain many puzzles and unsolved questions such as how brain regions interact to produce decisions and the roles of other neurotransmitters beyond DA. Although only monkey and human fMRI studies are reviewed here, other neuroimaging techniques such as EEG, MEG, and DTI can provide rich information on our brain systems and findings using this techniques should be integrated together to provide a better understandingof decision making systems. Examining clinical populations and their neural systems may provide rare opportunities otherwise unavailable (e.g. mOFC lesion patients) and will also greatly inform our knowledge on healthy brains. Indeed decision neuroscience is an inherently interdisciplinary field and the outlook is bright for researchers in various areas who seek the rewards of studying a new and exciting field.

References

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful Faces Have Variable Reward Value fMRI and Behavioral Evidence. *Neuron*, *32*(3), 537—551.

Ahn, W. Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A Model-Based fMRI Analysis with Hierarchical Bayesian Parameter Estimation. *Journal of Neuro-*

science, *Psychoogy*, *and Economics*, *4*(2), 95—110.

Arias-Carrion, O., & Popel, E. (2007). Dopamine, learning, and reward-seeking behavior. *Acta neurobiologiae experimentalis*, *67*(4).

Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., & Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology*, *94*(1), 327—337.

Belova, M. A., Paton, J. J., Morrison, S. E., & Salzman, C. D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron*, *55*, 970—984.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience*, *21*(8).

Berridge, K. C., & Aldridge, J. W. (2008). Special Review: Decision Utility, The Brain, and Pursuit of Hedonic Goals. *Social Cognition*, *26*(5), 621—646.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*(2), 619—639.

Brown, J. W., Bullock, D., & Grossberg, S. (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *Journal of Neuroscience*, *19*(23), 10502.

Busemeyer, J. R., Jessup, R. K., Johnson, J. G., & Townsend, J. T. (2006). Building bridges between neural models and complex decision making behaviour. *Neural Networks*, *19*(8), 1047—1058.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432—459.

Camps, M., Cortes, R., Gueye, B., Probst, A., & Palacios, J. M. (1989). Dopamine receptors in human brain: autoradiographic distribution of D2 sites. *Neuroscience*, *28*(2), 275—290.

Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology*, *363*, 615—641.

Carmichael, S. T., & Price, J. L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology*, *371*, 179—207.

Chiodo, L. A., Antelman, S. M., Caggiula, A. R., & Lineberry, C. G. (1980). Sensory stimuli alter the discharge rate of dopamine (DA) neurons: evidence for two functional types of DA cells in the substantia nigra. *Brain Research*, *189*(2).

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Current Opinion in Neurobiology*, *12*(2), 223—229.

Cohen, M. X. (2008). Neurocomputational mechanisms of reinforcement-guided learning in humans: A review. *Cognitive, Affective, and Behavioral Neuroscience*, *8*, 113—125.

Cortes, R., Camps, M., Gueye, B., Probst, A., & Palacios, J. M. (1989). Dopamine receptors in human brain: autoradiographic distribution of D1 and D2 sites in Parkinson syndrome of different etiology. *Brain Research*, *483*(1), 30—38.

Craig, A. D. (2009). A rat is not a monkey is not a human: comment on Mogil (Nature Rev. Neurosci. 10, 283—294 (2009)). *Nature Reviews Neuroscience*, 10.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, *319* (5867).

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*(2), 199—204.

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, *15*, 603—616.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical Substrates for Exploratory Decisions in Humans. *Nature*, *441*, 876—879.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, and Behavioral Neuroscience*, *8*, 429—453.

de Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C., & Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron*, *46*(4), 671—679.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*(1), 70—88.

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cognitive, Affective, and Behavioral Neuroscience*, *3*, 27—38.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, *84*, 3072—3077.

Dhar, R., Nowlis, S. M., & Sherman, S. J. (2000). Trying hard or hardly trying: An analysis of context effects in choice. *Journal of Consumer Psychology*, *9*(4), 189—200.

Ditterich, J., Mazurek, M. E., & Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature Neuroscience*, *6*(8), 891—898.

Dreher, J. C., Meyer-Lindenberg, A., Kohn, P., & Berman, K. F. (2008). Age-related changes in midbrain dopaminergic regulation of the human reward system. *Proceedings of*

the National Academy of Sciences, *105*(39).

Durstewitz, D., Kelc, M., & Gunturkun, O. (1999). A neurocomputational theory of the dopaminergic modulation of working memory functions. *Journal of Neuroscience*, *19*(7), 2807—2822.

Durstewitz, D., & Seamans, J. K. (2002). The computational role of dopamine D1 receptors in working memory. *Neural Networks*, *15*(4—6), 561—572.

Erk, S., Spitzer, M., Wunderlich, A. P., Galley, L., & Walter, H. (2002). Cultural objects modulate reward circuitry. *Neuroreport*, *13*(18).

Filoteo, J. V., Maddox, W. T., Simmons, A. N., Ing, A. D., Cagigas, X. E., Matthews, S., et al. (2005). Cortical and subcortical brain regions involved in rule-based category learning. *Neuroreport*, *16*, 111—115.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*, 1898—1902.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a Decision: Striato—Orbitofrontal Interactions in Reinforcement Learning, Decision Making, and Reversal. *Psychological Review*, *113*, 300—326.

Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience*, *120*, 497—517.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*, 495—501.

Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, *404*, 390—394.

Gold, J. I., & Shadlen, M. N. (2001). Neural Computations that Underlie Decisions About Sensory Stimuli. *Trends in Cognitive Science*, *5*, 10—16.

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, *36*, 299—308.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535—574.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*, *301*, 1104—1107.

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience*, *41*(1).

Guarraci, F. A., & Kapp, B. S. (1999). An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential Pavlovian fear conditioning in the awake rabbit. *Behavioural brain research*, *99*(2), 169—179.

Hall, H., Sedvall, G., Magnusson, O., Kopp, J., Halldin, C., & Farde, L. (1994). Distribution of D1-and D2-dopamine receptors, and dopamine and its metabolites in the human brain. *Neuropsychopharmacology*, *11*(4), 245—256.

Hampton, A. N., Adolphs, R., Tyszka, M. J., & O'Doherty, J. P. (2007). Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron*, *55*(4), 545—555.

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *Journal of Neuroscience*, *26*(32), 8360—8367.

Hikosaka, O., Sakamoto, M., & Usui, S. (1989a). Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *Journal of Neurophysiology*, *61*, 780—798.

Hikosaka, O., Sakamoto, M., & Usui, S. (1989b). Functional properties of monkey caudate neurons. II. Visual and auditory responses. *Journal of Neurophysiology*, *61*, 799—813.

Hikosaka, O., Sakamoto, M., & Usui, S. (1989c). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, *61*, 814—832.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *The Journal of Consumer Research*, *9*(1), 90—98.

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. The Journal of Consumer Research, 10(1), 31—44.

Johnson, J. G., & Busemeyer, J. R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychological Review*, *112*(4), 841—861.

Kahn, I., Yeshurun, Y., Rotshtein, P., Fried, I., Ben-Bashat, D., & Hendler, T. (2002). The role of the amygdala in signaling prospective outcome of choice. *Neuron*, *33*(6), 983—994.

Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLOS Biology*, *4*(8), e233.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, *21*.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*, 3683—3687.

Knutson, B., & Gibbs, S. E. B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology*, *191*(3), 813—822.

Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, *47*, 763—770.

Leon, M. I., & Shadlen, M. N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron*, *24*, 415—425.

Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology*, *67*(1), 145—163.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*, 150—157.

Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, *447*(7148), 1111—1115.

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837—841.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339—346.

McClure, S. M., Li, J., Tomlin, D., Cypert, K., Montague, L., & Montague, P. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, *44*(2), 379—387.

Mirenowicz, J., & Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*, *379*(6564), 449—451.

Mobbs, D., Greicius, M. D., Abdel-Azim, E., Menon, V., & Reiss, A. L. (2003). Humor modulates the mesolimbic reward centers. *Neuron*, *40*(5), 1041—1048.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936—1947.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., & Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, *43*(1), 133—143.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139—154.

O'Doherty, J. P. (2003). Can't learn without you: predictive value coding in orbitofrontal cortex requires the basolateral amygdala. *Neuron*, *39*(5), 731—733.

O'Doherty, J. P. (2007). Lights, camembert, action! The role of human orbitofrontal cortex in encoding stimuli, rewards, and choices. *Annals of the New York Academy of Sciences*,

1121, 254—272.

O'Doherty, J. P., & Bossaerts, P. (2008). Toward a Mechanistic Understanding of Human Decision Making: Contributions of Functional Neuroimaging. *Current Directions in Psychological Science*, *17*(2), 119—123.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329—337.

O'Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452—454.

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, *1104*, 35—53.

O'Doherty, J. P., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*(1), 95—102.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*, 223—226.

Paton, J. J., Belova, M. A., Morrison, S. E., & Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*, *439*, 865—870.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106).

Peyron, R., Garcia-Larrea, L., Gregoire, M. C., Costes, N., Convers, P., Lavenne, F., et al. (1999). Haemodynamic brain responses to acute pain in humans: sensory and attentional networks. *Brain*, *122*, 1765—1780.

Peyron, R., Laurent, B., & Garcia-Larrea, L. (2000). Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiologie Clinique*, *30*(5), 263—288.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, *105*(3), 1050.

Platt, M. L. (2002). Neural correlates of decisions. *Current Opinion in Neurobiology*, *12*, 141—148.

Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision

making under uncertainty. *Nature Neuroscience*, *11*(4), 398—403.

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*, 381—390.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545—556.

Ratcliff, R., & Smith, P. L. (2004). A Comparison of Sequential Sampling Models for Two—choice Reaction Time. *Psychological Review*, *111*, 333—367.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: a dynamic connectionist model of decision making. *Psychological Review*, *108*(2), 370—392.

Romo, R., & Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology*, *63*(3), 592—606.

Rorie, A. E., Gao, J., McClelland, J. L., & Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS ONE*, *5*(2), e9308.

Rosenkilde, C. E., Bauer, R. H., & Fuster, J. M. (1981). Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Research*, *209*(2).

Salas, R., Baldwin, P., de Biasi, M., & Montague, P. R. (2010). BOLD Responses to Negative Reward Prediction Errors in Human Habenula. *Frontiers in Human Neuroscience*, *4*, 36.

Schall, J. D. (2003). Neural correlates of decision processes: neural and mental chronometry. *Current Opinion in Neurobiology*, *13*(2), 182—186.

Schonberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, *27*(47), 12860.

Schonberg, T., O'Doherty, J. P., Joel, D., Inzelberg, R., Segev, Y., & Daw, N. D. (2010). Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson's disease patients: evidence from a model-based fMRI study. *Neuroimage*, *49*(1), 772—781.

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annual Review of Neuroscience*, *30*, 259—288.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, *13*(3), 900—913.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction

and reward. *Science*, *275*(5306), 1593—1599.

Schultz, W., & Romo, R. (1987). Responses of nigrostriatal dopamine neurons to high-intensity somatosensory stimulation in the anesthetized monkey. *Journal of Neurophysiology*, *57*(1), 201—217.

Seeman, P., Bzowej, N. H., Guan, H. C., Bergeron, C., Becker, L. E., Reynolds, G. P., et al. (1987). Human brain dopamine receptors in children and aging adults. *Synapse*, *1*(5).

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience*, *27*(18), 4826—4831.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. The *Journal of Consumer Research*, *16*(2), 158—174.

Sutton, R., & Barto, A. (1998). *Reinforcement learning*: MIT Press.

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differently recruits cortico-basal ganglia loops. *Nature Neuroscience*, *7*(8), 887—893.

Thorpe, S. J., Rolls, E. T., & Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Experimental Brain Research*, *49*(1), 93—115.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642—1645.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515—518.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, *41*, 281—292.

Tversky, A. (1972). Elimination by aspects: a theory of choice. Psychological *Review*, *79*(4), 281—289.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulaive Representations of Uncertainty. *Journal of Risk and Uncertainty*, *5*, 297—323.

Tversky, A., & Sattath, S. (1979). Preference trees. Psychological *Review*, *86*(6), 542—573.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*(10), 1179—1189.

Ungless, M. A., Magill, P. J., & Bolam, J. P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science*, *303*(5666), 2040—2042.

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson,

R. J., et al. (2004). Placebo-induced changes in FMRI in the anticipation and experience of pain. *Science*, *303*, 1162—1167.

Wallis, J. D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annual Review of Neuroscience*, *30*, 31—56.

Wang, G. J., Volkow, N. D., Logan, J., Fowler, J. S., Schlyer, D., MacGregor, R. R., et al. (1995). Evaluation of age-related changes in serotonin 5-HT2 and dopamine D2 receptor availability in healthy human subjects. *Life Sciences*, *56*(14).

Wang, X. J. (2008). Decision Making in Recurrent Neuronal Circuits. *Neuron*, *60*, 215—234.

Wang, Y., Chan, G. L. Y., Holden, J. E., Dobko, T., Mak, E., Schulzer, M., et al. (1998). Age-dependent decline of dopamine D1 receptors in human brain: a PET study. *Synapse*, *30*(1).

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, *5*(6), 483—494.

Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D. F., & Buchel, C. (2006). Dissociable systems for gain-and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience*, *26*(37).

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*(6), 464—476.