Neural and neurochemical basis of reinforcement-guided decision making

Abbas Khani and Gregor Rainer
Visual Cognition Laboratory, Department of Medicine, University of Fribourg, Switzerland

Submitted 15 December 2015; accepted in final form 24 May 2016

WHEN WE FACE SEVERAL OPTIONS, how do we choose the one that we choose? For example, how and why do we choose to buy a vanilla or a chocolate ice cream? How do we decide to travel to “Nice” over “Barcelona” for our vacation? How do we choose to work in academia or in business? We have to constantly decide to take a certain course of action over other competing and usually viable options. The kind of decisions we make range from very simple everyday decisions to decisions that are momentous and potentially have lifetime consequences. There is a multitude of decision variables that influence and guide our choices. These variables include, but are not limited to, current or future goals, subjective palatability, costs that are associated with the choice, quantity and quality, constraints like rules, and probability of success. Despite a recent surge in the number of studies investigating mechanisms and characteristics of decision making processes, the neural underpinnings of decision making are still emerging and not well understood. While we make a majority of our decisions in an effortless manner, patients with lesions in brain areas involved in decision processes as well as patients with a number of neuropsychiatric disorders make suboptimal decisions in many situations. Understanding the neural mechanisms of decision making will also facilitate development of novel therapeutic strategies and substances for such patients.

We aim in this article to review and organize recent trends about different aspects of decision making from a neuroscience perspective. In the last decade we have witnessed a surge in experimental research and computational modeling studying a multitude of variables pertaining to decision making. Reflecting the increased amount of experimental research, several reviews have focused on the role of specific brain regions or neurotransmitter systems in decision making, on certain aspects of decision making such as processing cost information or value representation, or in some cases on the type of scientific data such as neurophysiological data or data acquired in certain animal models. Rather than proposing another focused review of one aspect of decision making, we aim here to present a broad picture of the current understanding of decision making as a whole along with narrowly focused newer subsections like cannabinoids and exploratory decision making. Covering all aspects of a topic as broad as decision making in a single paper is challenging. Even so, we attempted to organize the present review in a clear layout and to integrate common themes from human and animal studies and from different approaches or methodologies in each subsection. Despite its broad coverage of key topics in reinforcement-based decision making, admittedly, this review does not encompass all pertinent ideas. We consider consensus opinion on most topics, although we indicate whenever there is a notable
controversy. After introducing perceptual decision making and reinforcer-guided decision making as two main approaches to decision making in cognitive neuroscience, we briefly discuss current theoretical understanding and models of reinforcement-guided decision making. In the next section, we evaluate cortical and subcortical neural circuitry involved in processing decision making. We then review major neurotransmitter systems that are key players in the processing of decision making. Finally, we survey scientific progress in the study of a relatively recent trend, i.e., exploratory decision making and foraging.

Reinforcement-Guided Decision Making

Depending on the approaches and applications, decision making has been categorized in different ways, although they include considerable overlapping elements and concepts. In cognitive neuroscience, perceptual decision making and reinforcement-guided decision making comprise two widely investigated forms of decision making (Botvinick 2012; Fetsch et al. 2014; Forstmann et al. 2010; Huk and Shadlen 2005; Kurther-Nelson and Redish 2010; Lee et al. 2012; Lee and Seo 2007; Morgan 1989; Newsome et al. 1989; Philiaistides and Sajda 2006; Ratcliff and Frank 2012; Ratcliff and Smith 2010; Rushworth et al. 2011, 2012; Sajda et al. 2009; Wendelken et al. 2009). In perceptual decision making, the perception of sensory stimuli has a key role in providing evidence in favor of or against a certain hypothesis concerning the state of the stimuli (Gold and Shadlen 2007). Motion direction discrimination from moving dots or a random-dot motion paradigm is an example of a widely used task in studies of perceptual decision making. In this task, a subset of dots (percent motion coherence) within a circle appears in an appropriate spatial displacement after a brief time lag to resemble apparent motion and the rest of the dots appear at random locations (Seidemann et al. 1998; Shadlen and Newsome 2001). In this paradigm the lower the percent motion coherence is, the more difficult the direction discrimination task is. On the other hand, reinforcement (or reward)-guided decision making models are based on economic (Von Neumann and Morgenstern 1944) and reinforcement learning (Sutton and Barto 1998) theories, and the focus is on the maximization of the value gained over a specific timescale. A closely related concept is value-based decision making, in which decisions are made primarily based on the subjective value associated with each alternative (Daw et al. 2006; Gold and Shadlen 2007; Wunderlich et al. 2010). Although we mainly focus on the influence of subjective value on choice behavior, it must be noted that, in addition to the subjective value associated with each alternative, decision making is also affected by objective perceptual properties such as salience of the choice-related stimulus (Lou et al. 2015).

Economic theories of decision making usually attempt to assign a value on a single scale to available actions or outcomes so that the decision of the subject is explained as the selection of the action with the maximum utility among all available alternatives (Glimcher 2002; Lee et al. 2012). A major drawback of models based on utilities is that they largely disregard constraints imposed by evolution as well as individual experiences. Similar to economic theories of decision making, the objective of reinforcement learning is to maximize future rewards. However, it must be noted that it rather takes into account experience-based estimates of a temporally discounted weighted sum of all future rewards in a specific timescale, which is often referred to as “return” (Lee and Seo 2007). Reinforcement learning is based on the premise that the prediction of an action’s value requires extensive knowledge, which is often not available, about the environment and consequences of performing that action. The actor-critic model, a recent model of the reinforcement learning theory, has suggested involvement of two distinct components: a “critic” that uses prediction error information to predict future rewards and an “actor” that uses the same information to update action values to ensure the choice of better ones (O’Doherty et al. 2004; Sutton and Barto 1998). An example of such distinction has experimentally been demonstrated in striatum, suggesting “critic” and “actor” roles for the ventral striatum and dorsal striatum, respectively (Kahnt et al. 2009; O’Doherty et al. 2004). Reinforcement learning theory explains how the experience of the animal modifies its value functions for a certain set of actions in a certain set of circumstances (state) and therefore influences its future choices. Schematic representations of these two models are depicted in Fig. 1. It should be noted that there are two different but closely related types of value functions in reinforcement learning theory. The action value function refers to the expected “return” from selecting a particular action in a particular state of the environment and is usually represented as $Q(s, a)$, in which $s$ and $a$ refer to the state of the environment and the action of the animal, respectively. The other value function, usually called the state value function and denoted by $V(s)$, refers to the sum of rewards expected from a certain state of the environment. The state value function is equal to the weighted (by the probability that the animal may take each action) average of action value functions available in a particular state (Lee et al. 2012). For example, consider an animal that is in a T-maze with equal probability to choose either the left or right arm of the maze and there are two pellets in the left arm and four pellets in the right arm. The state value function for the animal in this T-maze is equal to

$$2 \times 0.5 + 4 \times 0.5 = 3$$

The action value functions are equal to 2 and 4 for the actions of choosing the left and right arms, respectively. If the probability to choose changes to 20% vs. 80% for the left and right arms, respectively, then while the action value functions will remain the same as before for each action, the state value function will change for this state of the animal in the T-maze:

$$2 \times 0.2 + 4 \times 0.8 = 3.6$$

Several computational models have tried to explain reinforcement-guided decision making (Botvinick and Weinstein 2014; Doya 2008; Lee 2008). Recent studies have attempted to capture neural circuits of reinforcement-based decision making by adopting a multilevel approach (Soltani and Wang 2008; Wang 2002). In these studies (Fusi et al. 2007; Soltani et al. 2006; Soltani and Wang 2006), “a stochastic reward-dependent Hebbian plasticity rule was incorporated in a biophysically-based recurrent (attractor) network model of decision-making” (Soltani and Wang 2008). In this model, the learning rule is a Hebbian rule that is gated by reward signal and leads to long-term potentiation (LTP) if the choice of the network is rewarded or to long-term depression (LTD) if the choice is not
rewarded. This model captures the value in the form of “return” over several trials rather than the value of a single trial. Therefore, it is able to account for the leaky integration of reward value through mechanisms of LTP and LTD. The “return” for each option is the result of the weighted sum of updates in the past trials, with recent outcomes having stronger influence.

Recent work has brought together decision making studies by behavioral ecologists and neurobiologists. Behavioral ecologists have studied decision making by applying economic theory to explain ultimate causes of human and animal behavior. On the other hand, neurobiologists have studied the proximal causes of the same behavior by investigating how the brain produces specific behaviors (Glimcher 2002). Reconciling these two parallel approaches has suggested an evolutionary pressure on the brain’s decision making and learning systems to optimize the behavior in a manner that corresponds to behavioral ecologists’ description of the ultimate cause of behavior such as the need to secure food and achieve other kinds of rewards (Glimcher 2002; Rushworth et al. 2011).

Neural Circuitry Involved in Decision Making

Decision making is a complex process that involves integration of diverse information concerning past experiences, current needs and future goals, multimodal sensory inputs, internal state and emotional situation, as well as costs and risks (Felows 2004; Khani 2014). Therefore, it is not surprising that recent neurobiological studies have implicated a wide range of cortical and subcortical brain regions in decision making in one way or another. The fact that many forms of decision making are directed to receive different forms of reward implies that brain regions such as orbitofrontal cortex (OFC), ventral tegmental area (VTA), ventral and dorsal striatum, and amygdala are involved in reward expectation during decision making. Indeed, a recent study has demonstrated that the activity within a broader range of brain regions was modulated by the reward expectation (Ramayya et al. 2015). Since decision making is an executive process, the prefrontal cortex (PFC), which is responsible for governing executive functions, plays a pivotal role during decision making. Other brain regions are also involved in decision making, primarily through their connections to aforementioned brain regions providing other information such as sensory input or receiving motor commands. In this section, these different brain regions and circuitries are briefly introduced. Figure 2 shows some of the key circuitry governing the process of decision making.

Subcortical areas. Nucleus accumbens (NAc) is a key area involved in the processing of reward. Numerous studies have...
pointed out the importance of NAc and other subregions of striatum in reward processing during decision making and have demonstrated the existence of a strong correlation between the activity in the NAc and the expected value of outcomes critical for motivating behavior in the pursuit of reward (Burton et al. 2015; Mell et al. 2009; Patel et al. 2012; Stott and Redish 2014; Sugam et al. 2014). Several lines of study have also suggested a more specific role for subregions of striatum in processing different parameters of decision making. The most prominent role has been the encoding of reward representation and expectation (Burton et al. 2015; O’Doherty 2004; Schultz 2004), but other specific roles have also been suggested. For example, a number of studies have shown that NAc and other striatal subregions are involved in the processing of different cost information (Fig. 3) including effort (Ghods-Sharifi andFloresco 2010; Kurniawan et al. 2010; Mai et al. 2012; Schoupe et al. 2014), delay (Cardinal et al. 2001; Day et al. 2010; Hariri et al. 2006), and risk (Samanez-Larkin et al. 2010; Stopper andFloresco 2011; Sugam et al. 2014) during cost-benefit decision making tasks. The representation of different cost information in the NAc is dependent on a network that involves different regions of the PFC. Disconnection (Hauber and Sommer 2009) and functional magnetic resonance imaging (fMRI) (Croxson et al. 2009) studies suggested that a prefrontostriatal circuitry encompassing the anterior cingulate cortex (ACC) and NAc regulates effort-based decision making, whereas other studies demonstrated a critical role for the involvement of the OFC-NAc connections in regulating delay-based (Bezzina et al. 2008) and risk-based (Jung et al. 2010) decision making. On the other hand, dorsal striatum is involved in representing action value (Seo et al. 2012) and contributes to action selection and initiation (Hiebert et al. 2014) during decision making through integrating sensorimotor, motivational, and cognitive information within a network involving corticostriatal circuits (Balleine et al. 2007). It has also been suggested that the functional connectivity between substantia nigra and the dorsal striatum guides future choices by modulating action values as a function of both positive and negative reinforcements during past choices (Kahnt et al. 2009). Different subregions of striatum receive connections from the different areas of PFC and ACC (Berendse et al. 1992), which may suggest that topographically different areas of striatum are parts of separate fronto-striatal circuitries that process different aspects of decision making parameters (Rudebeck et al. 2006). Indeed, it has been proposed that as one moves from ventral and medial parts to dorsal and lateral parts of striatum, there is a shift from more prominent value encoding function to activity closely related to associative and motor aspects of decision making (Burton et al. 2015). Striatum is part of a larger complex of basal ganglia that are primarily involved in the coordination of movement. Some recent work has shown that the role of basal ganglia in decision making is not limited to striatum and that other areas including ventral pallidum (Ito and Doya 2009), substantia nigra pars reticulata (Hikosaka et al. 2006), and subthalamic nucleus (STN; Baunez and Lardeux 2011) as well as closely connected lateral habenula (Baker et al. 2015) within epithalamus are involved in the processing of decision making.

The STN is of particular interest since it is the main target for deep brain stimulation (DBS) in patients suffering from severe Parkinson’s disease (PD) that is intractable to pharmacotherapy alone (Benabid et al. 1994, 2009; Kumar et al. 1998; Limousin et al. 1995; Okun 2012), but this treatment induces some adverse effects including increased impulsivity (Frank et al. 2007; Robert et al. 2009) in a subset of patients. The increase in impulsive behavior is thought to result from the inactivation of the STN caused by DBS, such that the STN can no longer act as a “brake” to slow down decision making, particularly during high-conflict situations (Cavanagh et al. 2011; Frank et al. 2007). This means that in the absence of the STN subjects cannot wait to accumulate enough evidence to make decisions during high-conflict conditions. The connections between the medial PFC (mPFC) and the STN are particularly critical for regulating this function. Zavala et al. (2014) showed an increase in the theta-delta activity recorded from DBS electrodes implanted in the STN of PD patients. Using simultaneous midline frontal EEG recordings, they also demonstrated an increased theta-delta band coherence between two regions during high-conflict trials such that the activity in the midline frontal cortex was Granger causal to that in the STN, supporting the hypothesis that the brain uses frequency-specific communication channels to relay behaviorally important information (Zavala et al. 2014). In another study, Herz et al. (2016) demonstrated that low-frequency oscillations within the STN and corresponding mPFC-STN coupling predict how much evidence subjects accumulate before making a decision, such that subjects with stronger phase alignment in theta-delta frequency range between two regions have higher decision thresholds. However, studies investigating different aspects of

Fig. 3. Schematic representation of example cost-benefit tasks that are used in rats. There are different variations of these tasks and other paradigms that are used in rodents and other species to test cost-benefit reinforcement-based decision making. A: example effort-based decision making task. The animals are able to choose between a low and a high reward. However, there is a cost in the form of physical effort (climbing the barrier) to attain the high reward. B: delay-based T-maze decision making task. In this task, the rats can choose between a small and a large reward. However, the small reward can be attained immediately, but the rats have to wait during a delay period to reach the large reward. C: example risk-based decision making task. The rats initiate a trial by poking their nose into the hole, which makes 2 levers available for pressing. The animals can choose between pressing a lever that gives a small (1 pellet) but safe (in this example 100%) reward and pressing another lever that gives either a large reward (4 pellets) or no reward.
impulsivity have demonstrated that the role of STN in regulation of impulsivity is complex. While STN lesions increase impulsive action, resulting, for example, in premature responses during reaction time paradigms (Baunez et al. 1995; Uslaner and Robinson 2006), STN lesions decrease impulsive choice during delay-discounting decision tasks, such that animals are willing to wait longer for a large reward (Uslaner and Robinson 2006; Winstanley et al. 2005). This has been understood as a result of the involvement of the STN in associative and limbic circuits in addition to the traditional motor circuit (Baunez and Lardeux 2011), which prompted the suggestion that an enhancement in the incentive value of the reward underlies the decrease in impulsive choice following STN lesions, such that the motivation for the large reward outweighs impulsivity (Baunez et al. 2002; Bezzina et al. 2009).

Finally, it must be noted that some other subcortical areas including amygdala and VTA are also involved in certain aspects of decision making. VTA plays an essential role in decision making through its connections to the NAc and PFC that constitute mesolimbic and mesocortical dopaminergic pathways and are discussed in *Dopamine*. Basolateral amygdala (BLA) is part of an interconnected circuitry that also includes the OFC and NAc and represents reward expectations and associations (Saddoris et al. 2005; Schoenbaum and Roesch 2005). With amygdala-OFC disconnection by unilateral lesions of these two structures in opposing hemispheres in monkey, it has been demonstrated that amygdala-OFC interaction is crucial for the adjustments in decision making following changes in the outcome (Baxter et al. 2000). Similarly, an fMRI study of two rare patients with bilateral lesions in the amygdala showed a profound change in the ventromedial PFC (vmPFC) corresponding to reward expectation and choice (Hampton et al. 2007). These studies support an important role for the amygdala in establishing reward value representation in the OFC/vmPFC. Studies have shown that amygdala is important for emotional and autonomic aspects of decision making, and patients with damage to amygdala were unable both to perform optimal decision making and to produce anticipatory skin conductance responses (Bechara et al. 1995; Gupta et al. 2017). The role of amygdala in emotional aspects of the decision making has been reflected in differential amygdala-vmPFC connectivity depending on the emotional input in the decision task. In decisions concerning moral judgments, the amygdala-vmPFC connectivity was the highest for pure emotional assessments and the lowest for pure utilitarian assessments, which supports the hypothesis that amygdala provides affective value of the choice to the vmPFC (Shenhav and Greene 2014). An interesting finding was an interaction between sex and laterality of amygdala function such that men with unilateral damage to the right amygdala showed more deficits in decision making and social behavior whereas damage to the left amygdala was more deleterious to women patients (Gupta et al. 2011). Several other single-neuron studies in humans (Jenison et al. 2011) and nonhuman primates (Grabenhorst et al. 2012) as well as fMRI studies in human subjects (Ousdal et al. 2014) have indicated that amygdala encodes the value of the expected reward or certain aspects of reward value such as reward magnitude (Bermudez and Schultz 2010) and reward timing (Bermudez et al. 2012). More specifically, it has been shown that not only is amygdala involved in early stages of decision making in encoding benefits and values but it also predicts economic choice (Grabenhorst et al. 2012) and plays a key role in processing different costs such as physical effort (Floresco and Ghods-Sharifi 2007; Ghods-Sharifi et al. 2009), delay (Winstanley et al. 2004), and risk (Bechara et al. 1999; Ghods-Sharifi et al. 2009; Zeeb and Winstanley 2011). The role of amygdala in regulating effort-based decision making is, at least partially, mediated by amygdala-prefrontal circuitry including the BLA and ACC (Floresco and Ghods-Sharifi 2007). Similarly, it has been shown that the OFC-BLA pathway is important for regulating flexibility to change behavior during decision making whereas the mPFC-BLA pathway is necessary for regulating delay-based decisions or decreased impulsivity (Churchwell et al. 2009). Regarding the risk-based decisions, the role of amygdala has been linked closely to the circuits including the BLA, PFC, and striatum. Zeeb and Winstanley (2013) demonstrated that functional disconnection of the OFC and BLA impairs acquisition of a rat gambling task. Additionally, genotyping cathecol-O-methyltransferase (COMT) polymorphism combined with resting-state functional connectivity showed that the Met allele carriers of COMT had a higher propensity to take risk paralleled with a decreased resting-state functional connectivity between the OFC and amygdala (Gao et al. 2016). mPFC-BLA connections also play a role in regulating decisions that might be risky. Taking the advantage of separate axonal pathways from the BLA to the mPFC and vice versa, St Onge et al. (2012) selectively disrupted bottom-up and top-down connections between the mPFC and BLA. The authors demonstrated that top-down, but not bottom-up, communication between these two regions is necessary for controlling risky decisions. They also showed that BLA-NAc circuitry promotes taking risky decisions.

**Prefrontal and cingulate areas.** There is consensus that PFC plays a critical role in decision making, with evidence ranging from an extensive neurological history of decision making disorders in patients with frontal lesions to findings of decision making-related subprocesses in PFC (Fellows 2004; Kennerley and Walton 2011; Krawczyk 2002). Recently there has been a surge in research studying the functional specialization of decision making in subregions of PFC. Depending on the level of the specialization under investigation, techniques of the study, and the species used in the study, different functional dissociations among different levels of frontal subregions have been suggested. Rushworth et al. (2011) have suggested that there are at least four frontal subregions with distinct roles in reinforcement-guided decision making. These areas are as follows: “ventromedial prefrontal cortex and adjacent medial orbitofrontal cortex (vmPFC/mOFC), lateral orbitofrontal cortex (IOFC), ACC, and a lateral anterior PFC (aPFC) region in, or at least adjacent to, the lateral part of the frontal pole.” We review here the distinct roles that these regions play during decision making, and, when available, we briefly introduce more recent work revealing finer and detailed functional specialization of frontal subregions during decision making. It is noteworthy that many studies have investigated the role of OFC in certain aspects of decision making as a whole and have not differentiated between lateral and medial parts of these studies, and we will use “OFC” rather than “mOFC” or “IOFC” when we refer to these studies.

Numerous studies have shown a partial functional dissociation and partial functional overlap between the ACC and OFC in various facets of decision making (Khani 2014; Lee et al. 2016). mPFC-BLA connections also play a role in regulating delay-based decisions or decreased impulsivity (Floresco and Ghods-Sharifi 2007). The role of amygdala in regulating effort-based decision making is, at least partially, mediated by amygdala-prefrontal circuitry including the BLA and ACC (Floresco and Ghods-Sharifi 2007). Similarly, it has been shown that the OFC-BLA pathway is important for regulating flexibility to change behavior during decision making whereas the mPFC-BLA pathway is necessary for regulating delay-based decisions or decreased impulsivity (Churchwell et al. 2009). Regarding the risk-based decisions, the role of amygdala has been linked closely to the circuits including the BLA, PFC, and striatum. Zeeb and Winstanley (2013) demonstrated that functional disconnection of the OFC and BLA impairs acquisition of a rat gambling task. Additionally, genotyping cathecol-O-methyltransferase (COMT) polymorphism combined with resting-state functional connectivity showed that the Met allele carriers of COMT had a higher propensity to take risk paralleled with a decreased resting-state functional connectivity between the OFC and amygdala (Gao et al. 2016). mPFC-BLA connections also play a role in regulating decisions that might be risky. Taking the advantage of separate axonal pathways from the BLA to the mPFC and vice versa, St Onge et al. (2012) selectively disrupted bottom-up and top-down connections between the mPFC and BLA. The authors demonstrated that top-down, but not bottom-up, communication between these two regions is necessary for controlling risky decisions. They also showed that BLA-NAc circuitry promotes taking risky decisions.

**Prefrontal and cingulate areas.** There is consensus that PFC plays a critical role in decision making, with evidence ranging from an extensive neurological history of decision making disorders in patients with frontal lesions to findings of decision making-related subprocesses in PFC (Fellows 2004; Kennerley and Walton 2011; Krawczyk 2002). Recently there has been a surge in research studying the functional specialization of decision making in subregions of PFC. Depending on the level of the specialization under investigation, techniques of the study, and the species used in the study, different functional dissociations among different levels of frontal subregions have been suggested. Rushworth et al. (2011) have suggested that there are at least four frontal subregions with distinct roles in reinforcement-guided decision making. These areas are as follows: “ventromedial prefrontal cortex and adjacent medial orbitofrontal cortex (vmPFC/mOFC), lateral orbitofrontal cortex (IOFC), ACC, and a lateral anterior PFC (aPFC) region in, or at least adjacent to, the lateral part of the frontal pole.” We review here the distinct roles that these regions play during decision making, and, when available, we briefly introduce more recent work revealing finer and detailed functional specialization of frontal subregions during decision making. It is noteworthy that many studies have investigated the role of OFC in certain aspects of decision making as a whole and have not differentiated between lateral and medial parts of these studies, and we will use “OFC” rather than “mOFC” or “IOFC” when we refer to these studies.

Numerous studies have shown a partial functional dissociation and partial functional overlap between the ACC and OFC in various facets of decision making (Khani 2014; Lee et al. 2016). mPFC-BLA connections also play a role in regulating delay-based decisions or decreased impulsivity (Floresco and Ghods-Sharifi 2007). The role of amygdala in regulating effort-based decision making is, at least partially, mediated by amygdala-prefrontal circuitry including the BLA and ACC (Floresco and Ghods-Sharifi 2007). Similarly, it has been shown that the OFC-BLA pathway is important for regulating flexibility to change behavior during decision making whereas the mPFC-BLA pathway is necessary for regulating delay-based decisions or decreased impulsivity (Churchwell et al. 2009). Regarding the risk-based decisions, the role of amygdala has been linked closely to the circuits including the BLA, PFC, and striatum. Zeeb and Winstanley (2013) demonstrated that functional disconnection of the OFC and BLA impairs acquisition of a rat gambling task. Additionally, genotyping cathecol-O-methyltransferase (COMT) polymorphism combined with resting-state functional connectivity showed that the Met allele carriers of COMT had a higher propensity to take risk paralleled with a decreased resting-state functional connectivity between the OFC and amygdala (Gao et al. 2016). mPFC-BLA connections also play a role in regulating decisions that might be risky. Taking the advantage of separate axonal pathways from the BLA to the mPFC and vice versa, St Onge et al. (2012) selectively disrupted bottom-up and top-down connections between the mPFC and BLA. The authors demonstrated that top-down, but not bottom-up, communication between these two regions is necessary for controlling risky decisions. They also showed that BLA-NAc circuitry promotes taking risky decisions.
Many studies have demonstrated a role for the OFC in value representation during reinforcement-guided decision making (Izquierdo et al. 2004; Padoa-Schioppa and Assad 2006; Roesch and Olson 2004; Wallis 2007). It has been suggested that the neurons within the OFC encode subjective value in a subjective scale independent from the available options in a way that it represents a common currency for the value of options (Levy and Glimcher 2012; Padoa-Schioppa and Assad 2008). Some other studies have also shown a role for the ACC in representing value signals (Hosokawa et al. 2013; Kennerley et al. 2009; Sescousse et al. 2010). Other studies have shown that the value representations in these two frontal regions are dissociable. For example, a lesion study in the macaque (Rudebeck et al. 2008b) and a study in human patients with focal frontal lobe damage (Camille et al. 2011) have shown that the OFC and ACC represent value in different contexts. While the OFC is important for optimal decision making that depends on stimulus-outcome associations, the ACC is essential for decision making that depends on action-outcome associations. A single-neuron recording study in nonhuman primates (Luk and Wallis 2013) showed that such a dissociation in representation of stimulus and action values during decision making is partial (Khani 2014). It has also been suggested that a subset of neurons within the ACC represent integrated value of options, taking into account different decision variables (Hosokawa et al. 2013; Kennerley et al. 2009, 2011). Furthermore, recent work (Burton et al. 2014; Metereau and Dreher 2014; Plassmann et al. 2010) has shown that most of the value representation in the OFC is indeed the function of medial parts of the OFC and ventromedial prefrontal areas (mOFC/vmPFC). In fact, recent studies showed distinct functions for the vmPFC together with adjacent mOFC compared with the IOFC. These studies suggest that iOFC is involved in value-credit assignment and reward-value learning independent from the decisions that animals make whereas vmPFC/mOFC does not have a role in value-credit assignment but plays a key role in evaluation, value-guided decision-making, and maintenance of a choice over successive decisions (Noonan et al. 2010b, 2012). More detailed functional specialization has been revealed within the vmPFC/mOFC and the IOFC. For example, Bouret and Richmond (2010) have demonstrated that the vmPFC and OFC in monkeys differed in encoding values driven by internal and external motivations such that the vmPFC neurons were more sensitive to internal factors like satiety and the OFC neurons were more sensitive to external factors such as visual cues. Another neuroimaging study in human subjects (Sescousse et al. 2010) showed that while several brain regions including the ACC, ventral striatum, anterior insula, and midbrain encode subjective value of rewards regardless of the reward type, subregions of the OFC represent different types of reward values. While the anterior lateral OFC processes monetary gains, the phylogenetically older posterior lateral OFC region processes more basic erotic stimuli, suggesting an increase in complexity along a postero-anterior axis in the IOFC to more abstract representations.

In addition to the gross value of a choice, the costs incurred to achieve that value have an important influence on animal and human calculations of the net value of a choice. In decision making paradigms, costs can be in the form of a time delay, during which the subject has to wait for an expected reward, or in the form of physical effort necessary to expend to gain an expected benefit (Khani et al. 2015). Several studies have shown that the ACC, but not the adjacent prelimbic and infralimbic areas, is responsible for the processing of effort cost (Walton et al. 2002, 2003; Walton et al.). Indeed, a recent lesion study in rats has shown that the ACC and OFC are involved in the processing of distinct decision costs during decision making (Rudebeck et al. 2006). While ACC-lesioned rats made decisions similar to control animals for delayed rewards, they exhibited pronounced deficits during decisions involving effort costs. OFC-lesioned rats exhibited an opposite pattern of disruption: they were highly impaired on decisions involving delay costs but performed just as well as control animals in decisions based on the effort costs. When an equal cost was necessary to achieve the low reward, all lesioned rats shifted to high-reward choices, confirming that these results were specific to differential cost-benefit decision making and were not confounded by the spatial or reward memory. A similar double dissociation between OFC and ACC in the processing of delay and effort decision costs has been demonstrated in a human neuroimaging experiment as well (Prevost et al. 2010). However, the dissociation in the processing of different costs is partial at the level of single neurons. Hosokawa et al. (2013) demonstrated that single neurons in the OFC are sensitive to delay-based decision tasks whereas a population of neurons in the ACC was modulated by both effort- and delay-based decisions. Regarding the role of the OFC in the processing delay costs, another group has demonstrated dissociable function between lOFC and mOFC in the processing of delay costs such that IOFC-lesioned rats exhibited a more impulsive pattern of choice compared with control animals while mOFC-lesioned rats exhibited an opposite effect showing an increase in the number of large delayed rewards (Mar et al. 2011). However, there is a need for caution in interpreting the results of this study (Mar et al. 2011) in conjunction with the former study (Rudebeck et al. 2006) since they used different decision making tasks and brought about opposing results when the entire OFC region was lesioned. Finally, we point out another important dissociation between the functions of the ACC and the OFC. While the OFC is involved in the processing of emotional aspects of decisions (Simmons et al. 2014; Sommer et al. 2009), the ACC plays a critical role in decisions that require valuation of social information or interaction with other individuals (Noonan et al. 2010a; Rudebeck et al. 2008a; Walton et al. 2007).

Frontopolar cortex (FPC) is the most anterior region of the frontal lobe. While the activity in the FPC during planning tasks has long been demonstrated (Owen et al. 1996), the functions of this area during decision making and cognitive control are emerging recently. It has also been suggested that the FPC, including lateral aPFC, which lies within or adjacent to the FPC, is involved in the integration of other executive functions. For example, a neuroimaging study showed an increased sustained activity in the FPC when the demand for temporal integration increased (Yarkoni et al. 2005). It has also been suggested that the FPC is preferentially active during exploratory decision making compared with value-based decision making (Daw et al. 2006) and self-generated decisions as opposed to decisions following trial-by-trial instructions (Tsuchimoto et al. 2010). More recent work has shed light on the role of the FPC during decision making, indicating a key role in...
evaluating alternative unchosen courses of action (Boorman et al. 2009, 2011; Bunge and Wendelken 2009). Boorman et al. (2009) have demonstrated that unlike the existence of a positive correlation between the value of chosen options and the activity in the vmPFC/mOFC, there is a negative correlation between the activity in the FPC and the value of chosen options. However, they showed a positive correlation between the activity in the FPC and the unchosen option. This is called the value of “counterfactual” choice (Boorman et al. 2011; Rushworth et al. 2011). The value of the unchosen option is important in adaptive behavior and in guiding future decisions of whether to keep on choosing the same option or switch to the unchosen option in new encounters. In fact, it has been shown that the individual differences in the activity of the FPC predict interindividual differences in effectively adapting behavior and switching to other options in forthcoming trials (Boorman et al. 2009).

**Neuromodulatory Systems Involved in Decision Making**

Brain structures that are discussed in Reinforcement-Guided Decision Making connect to each other as well as many other areas and make a complex network that processes information and facilitates cognitive control and optimal decision making. These connections use a variety of neurotransmitters and neuromodulators to effectively communicate and thereby process available information and generate appropriate and beneficial responses/behavior. Investigating the underlying neurochemistry of decision making processes is essential for our understanding of the neural mechanisms of decision making. In this section we briefly review the main neurotransmitter systems involved in decision making.

**Dopamine.** The involvement of the dopamine (DA) system in different aspects of decision making is perhaps the most investigated among neurotransmitter systems. However, the role of DA in decision making is not completely understood. Midbrain DA neurons encode a range of parameters related to the value of the rewards in the past and in the future (Schultz 2002). These reward signals range from motivation (Satoh et al. 2003), reward magnitude (Calaminus and Hauber 2009), valence (Koob 1996), and probability and uncertainty (Fiorillo et al. 2003) to reward memory (Yamagata et al. 2015) and are integrated with various other information regarding other aspects of the available behavioral options such as the costs and risks to evaluate representations of different options and focus the network on the preferred representation to proceed to action (Assadi et al. 2009). Interestingly, these neurons innervate key brain structures (PFC, amygdala, dorsal striatum, and NAc) implicated in economic decision making (Gan et al. 2010; Glimcher et al. 2005). The role of the DA system in decision making is partly attributed to its role in predicting reward and associated prediction errors. Some DA neurons encode the discrepancy between the reward value and its expected value [i.e., reward prediction error (RPE)] rather than the absolute value of the reward (Schultz 1998; Tobler et al. 2005). RPE can be positive when there is an unexpected reward, zero when an expected reward is obtained, or negative when an expected reward is omitted. Seminal work by Schultz and colleagues (Schultz 1997; Schultz et al. 1997) demonstrated that mesolimbic DA neurons encode cues that predict reward and respond to errors in those predictions. Phasic DA release in the NAc contributes to the processing of value-related parameters during decision making (Gan et al. 2010). The phasic DA release in the NAc core is involved in the learning stimulus-reward associations in which incentive salience is attributed to reward cues (Flagel et al. 2011), whereas phasic DA dynamics in the NAc shell tracks motivationally salient stimuli and encodes both predictive cues and rewards (Saddoris et al. 2015). The involvement of DA signaling in the processing of decision making has been captured in several computational models (Friston et al. 2014; McClure et al. 2003). Most of these models accounting for DA role during decision making center around reward-predictive coding by DA neurons (Egelson et al. 1998; McClure et al. 2003). Some of these models employ Bayesian inference methods and have shown that DA neurons encode the certainty about the expected outcome and help optimize decision making by assimilating RPEs (Friston et al. 2013, 2014; Schwartenbeck et al. 2015).

In addition to the involvement of the DA system in the valuation of expected outcomes, several studies have shown that DA signaling is necessary to overcome different forms of costs during cost-benefit decision making tasks (Assadi et al. 2009; Saddoris et al. 2014; Sugam et al. 2012). In particular, the mesolimbic dopaminergic projection system that innervates the NAc has been implicated in effort-based decision making (Day et al. 2010; Hauber and Sommer 2009; Mai et al. 2012; Salamone et al. 1994). Studies have also shown that DA signaling in the NAc encodes other types of decision cost such as delay (Day et al. 2010; Saddoris et al. 2014) and risk (Sugam et al. 2012, 2014). The involvement of the mesocortical dopaminergic pathway in decision making has been controversial and remains a subject of debate. For example, while one study failed to find evidence for a role of dopaminergic neurons projecting from the VTA to the ACC during effort-based form of cost-benefit decision making (Walton et al. 2005), another group have demonstrated that blockade or downregulation of dopaminergic neurotransmission in the ACC impairs effort-based decision making in rats such that they are less willing to put out effort to gain a higher amount of reward (Schweimer and Hauber 2005, 2006). As the latter authors discussed, the discrepancy between results may largely arise because of the differences in the extent of DA depletion, such that an extensive DA depletion or blockade of D1 receptors in the ACC is necessary to disrupt effort-based decision making. It has also been reported that blockade of DA receptors in the OFC increases impulsive choice (Zeeb et al. 2010), suggesting a role for the DA signaling in the OFC in delay-based decision making.

The role of the DA system in the processing of decision making has also been implicated in a number of neurological and psychiatric disorders. For example, it has been shown that DA is involved in temporal aspects of decision making such that PD patients “on” dopaminergic medication were better in “Go” learning whereas patients “off” dopaminergic medication were better in slowing down or “NoGo” learning (Moustafa et al. 2008). Consistent with this work, other studies (Osman et al. 2014; Torta et al. 2009) have shown that DA medication in PD patients shifts decisions to impulsive patterns of choice and suboptimal strategies. These effects are dose dependent (Torta et al. 2009) and support the “DA overdose hypothesis,” and they suggest that DA medication can lead to adverse effects in executive functions such as decision making in PD patients...
(Osman et al. 2014). Decline in the concentrations of DA neurotransmission has also been linked to impoverished stimulus-reward learning and thereby impaired reinforcement-guided decision making in older adults (Mell et al. 2009). Concurrently, another study demonstrated that l-DOPA (a DA precursor) restored deficits in RPE signal in the NAc and improved decision making in a population of old subjects (Chowdhury et al. 2013).

Serotonin. Similar to DA, serotonin has long been implicated in adaptive behavior and reinforcement learning, although it has been studied considerably less extensively than DA. It has been suggested that serotonin is involved in processing several decision making-related variables such as reward assessment, cost assessment, impulsivity, harm aversion, and anxious states (Asker et al. 2013). Notwithstanding the crucial role of DA in decision making, the DA system cannot solely secure optimal decision making in the face of perturbations in other neuropeptides involved in decision making. Importantly, it has been suggested that a balance in the functions of DA and serotonin neurons is necessary for proper adaptive behavior and optimal decision making (Eppinger et al. 2011; Whitaker-Azmitia et al. 1990). Human serotonin transporter promoter length polymorphism (5-HTTLPR) affects several psychiatric disorders. It has been shown that polymorphism of this gene influences the pattern of decision making such that subjects homozygous for the short allele of 5-HTTLPR choose more disadvantageously than subjects homozygous for the long allele of the gene. These changes in the pattern of decision making as a result of 5-HTTLPR have been reported among healthy volunteers (Homberg et al. 2008) as well as major depressive disorder patients (Must et al. 2007) and obsessive-compulsive disorder patients (da Rocha et al. 2008). Consistent with these findings, other authors (Morgan et al. 2006; Quednow et al. 2007) have shown that MDMA (“ecstasy”)-induced dysregulation of the serotonergic system impairs decision making and elevates impulsivity. Furthermore, depletion of serotonin leads to suboptimal decision making by impairing discrimination between magnitudes of expected gains associated with different choices (Rogers et al. 2003). Serotonergic neurons of the dorsal raphe nucleus, the main origin of serotoninergic neurons along with mediod raphe nucleus, project to diverse cortical and subcortical areas involved in decision making, and their own firing activity is under frontal feedback control (Homberg 2012). It has been demonstrated that single neurons within dorsal raphe nucleus encode reward parameters as well as specific sensorimotor events, including stimulus identity and response direction (Ranade and Mainen 2009). Serotonin has been implicated in the processing of both negative and positive outcome value. While Cools et al. (2008) argued that the decline in the levels of brain serotonin impairs decision making through enhancing the effects of punishment compared with reward, more recently Seymour et al. (2012) showed that serotonin depletion selectively impaired behavioral and neural representations of reward outcome value, thereby impairing effective comparison of rewards and punishments.

Serotonin has also been implicated in the processing of different costs and probability of outcomes. Unlike DA, it seems that serotonin is not involved in processing effort-based decision making (Denk et al. 2005). However, we caution that there is not much work investigating the possible role of serotonin in the processing of this type of decision costs. Indeed, a recent study raises the possibility that serotonin might be involved in effort-based decision making (Kosheleff et al. 2012). On the other hand, there is plenty of evidence suggesting a key role for serotonin in the processing of delay costs and impulsive patterns of choice (Bizot et al. 1999; Denk et al. 2005; Hadamitzky et al. 2009; Schweighofer et al. 2008; Wright et al. 2012). Importantly, it has been demonstrated that lesions of dorsal and medial raphe nuclei, which send serotoninergic projections to the frontal cortex, lead to a shift from choosing large delayed rewards to selecting a small immediately available reward (Mobini et al. 2000; Wogar et al. 1993). Consistent with these findings, a recent study showed that activation of serotoninergic neurons in the dorsal raphe nucleus is necessary for waiting for long-delayed rewards and suggests that enhanced serotonin signaling facilitates waiting behavior when forthcoming rewards are expected (Miyazaki et al. 2012). Finally, serotonin has been implicated in the processing of risk-based choice and choices under circumstances of uncertainty (Balasubramani et al. 2014; Ishii et al. 2015; Long et al. 2009; Rock et al. 2013). Along the same line, Koot et al. (2012) used rodent versions of the Iowa gambling task and the probabilistic delivery task and demonstrated that serotonin precursor depletion leads to poor decision making and increased gambling proneness.

Cannabinoids. Endocannabinoids have emerged as major modulators of synaptic activity acting principally in a retrograde manner. They are usually released from the postsynaptic neurons and travel back to the presynaptic terminals modulating the release of the neurotransmitters from those axonal ends (Castillo et al. 2012). Cannabinoid receptors are distributed throughout the brain (Kano et al. 2009) and therefore are involved in regulating a wide range of brain functions including feeding behavior (Engeli 2012), memory (Hampson et al. 2011; Jacob et al. 2012), stress responses (Ganon-Elazar and Akirav 2009), motor control (Dubreucq et al. 2013), and pain processing (Lee et al. 2013; Martin et al. 1999). Cannabinoid receptor type 1 (CB1R) is thought to mediate most of the brain functions of the cannabinoid system and is densely distributed in many brain areas involved in decision making, with a high abundance in the cingulate cortex and association cortical regions of the frontal lobe (Eggan and Lewis 2007; Filbey et al. 2010). It has also been shown that patients experiencing a first episode of schizophrenia who use cannabis have smaller anterior cingulate gray matter compared with patients who do not use cannabis or healthy control subjects (Szegasko et al. 2007). While these lines of evidence together with considerable evidence correlating the use of various cannabis derivatives with altered cognition and decision making suggest a modulatory role for the cannabinoid system during decision making, the role of cannabinoid signaling in decision making pathways has not received much attention to date.

Most of the evidence implicating the cannabinoid system in modulating decision processes comes from human studies in cannabis abusers or recreational consumers. These experiments generally associate cannabis use with increased impulsive pattern of choice, risky decision making, and gambling among patients and healthy volunteers (Fischer et al. 2015; Fridberg et al. 2010; Gonzalez et al. 2012; Leppink et al. 2014; Solowij et al. 2012). Such disruptions in patterns of decision making are not restricted to heavy cannabis users, and similar deficits have
been documented among recreational users (Griffith-Lendering et al. 2012; Moreno et al. 2012) as well as adolescents starting to use cannabis (De Bellis et al. 2013). Indeed, a recent neuroimaging study showed an increased activity in response to wins during a monetary decision making task in core areas associated with decision making and revealed that individual differences in decision making and reward processing predict future changes in cannabis use such that individuals with biased choices toward immediate rewards are more likely to increase drug use (Cousijn et al. 2013). There is also evidence suggesting sex-specific deficits among cannabis users, such that cannabis use was more consistently associated with poorer episodic memory performance in women than men and poorer decision-making performance in men but not women (Crane et al. 2013).

In addition to studies among cannabis users, recent work has started to delineate mechanisms by which the cannabinoid system modulates decision making and value processing. Imaging studies during decision making tasks showed that abstinent cannabis users recruit vmPFC/mOFC more strongly and IOFC less strongly compared with control subjects (Bolla et al. 2005; Vaidya et al. 2012). Another neuroimaging study in rodents has shown that cannabinoid agonist administration increases BOLD activity in several brain areas including the NAc and VTA that are involved in reward and value processing (Shah et al. 2004). Consistent with this study, Deshmukh and Sharma (2012) showed an increased food intake following infusion of a cannabinoid agonist into the NAc shell and suggested that the cannabinoid system promotes eating motivation by enhancing the incentive value of food. Other studies suggest that cannabinoid receptors within the NAc (Bloomfield et al. 2014) and VTA (Melis et al. 2012) influence the processing of reward value by modulating DA signaling (Hernandez and Cheer 2015). For example, a recent work demonstrated that the blockade of CB1Rs attenuates DA release in the NAc in response to rewarding medial forebrain bundle stimulation, thereby decreasing performance for the electrical reward (Trujillo-Pisanty et al. 2011). Previous work has also shown that the cannabinoid system is involved in cost-benefit decision making based on different costs. These studies have generally been conducted in cannabis users or have used systemic administration of cannabinoids and resulted in inconsistent results. Some studies demonstrated increased impulsive choice (Hernandez et al. 2014; Lofflin et al. 2014) after cannabinoid administration, while others showed unaffected or decreased impulsive choice (McDonald et al. 2003; Pattij et al. 2007; Wiskerke et al. 2011). Along the same lines, it has been proposed that genetic factors influence the effects of cannabinoids on impulsive pattern of choice (Bohmhower et al. 2013) and polymorphisms in CB1R gene are involved in differential trait impulsivity (Ehlers et al. 2007). On risk-based decision making, a computational modeling of decision making patterns suggested that cannabis abusers, compared with healthy volunteers, are underinfluenced by the magnitude of losses but are overinfluenced by gains (Fridberg et al. 2010). There is also work suggesting cannabinoid modulation of effort-based cost-benefit decision making. Sink et al. (2008) showed that CB1R blockade reduced the willingness of rats to press a lever to acquire preferred food. We have recently shown that local activation of the cannabinoid system in frontal regions modulates effort-based and delay-based decision making (Khani et al. 2015). Consistent with lesion studies (Rudebeck et al. 2006), these effects were double dissociable. Activation of the cannabinoid system in the ACC but not in the OFC induced a shift in the choice of rats such that they chose low-reward/low-effort choices more frequently compared with rats that received a drug-free vehicle. Similarly, rats performing a delay-based decision making task were less willing to wait for large delayed rewards after the activation of the cannabinoid system in the OFC but not in the ACC. Immunohistochemistry experiments suggested that these effects were caused by dysregulation in cannabinoid system modulation of GABAergic interneurons and to a lesser degree dopaminergic, serotonergic, and glutamatergic neurons (Khani et al. 2015).

Other neurotransmitter systems. Several other neurotransmitter systems play a role in some aspects of decision making. A recent study has corroborated the idea that moderate increases in the levels of catecholamines, but not cortisol, enhance decision making performance (Pabst et al. 2013). Usher and Dayalaa (2002) suggested that the noradrenergic system is involved in decision making through modulation of the synaptic efficiency of neural circuits by phasic responses in the brain nucleus locus coeruleus (LC). In line with this suggestion, Bouret and Sara (2005) argued that phasic activation of noradrenergic neurons of the LC reset network dynamics facilitating rapid behavioral adaptation to changing environmental imperatives. There is experimental evidence supporting the hypothesis that phasic norepinephrine transmission is involved in decision making. Namely, Clayton et al. (2004) demonstrated that phasic activation of LC noradrenergic neurons preceded both correct and incorrect behavioral responses in a forced-choice discrimination task but those neurons were not activated by stimuli that failed to elicit response. These authors suggested that the LC neurons regulate the behavioral outcome of decisional processes. Additionally, it has been suggested that P3, an important component of event-related brain potential during decision processes, reflects phasic activity of LC noradrenergic neurons in response to the outcome of internal decision making processes (for a review, see Nieuwenhuis et al. 2005). Several computational models have also implicated norepinephrine in the processing of decision making (Brown et al. 2005; Shea-Brown et al. 2008; Yu and Dayan 2005). For example, Yu and Dayan (2005) proposed that both acetylcholine and norepinephrine are involved in the computations of uncertainty in noisy environments. However, they suggested that these neuromodulators act in a different way. While acetylcholine signals expected uncertainty, norepinephrine signals unexpected uncertainty when sudden context switches produce strongly unexpected observations. Acetylcholine has long been implicated in apt cognitive functions. As mentioned above, it has been suggested to signal expected uncertainty when the unreliability of predictive cues within a context is known (Yu and Dayan 2005). More recently, several studies have shown specific contributions of the cholinergic system to different aspects of decision making particularly in processing cost-benefit decision making. It has been reported that acute administration of central muscarinic cholinergic receptor antagonists induces impulsive pattern of choice (Mendez et al. 2012). Another study has also implicated muscarinic cholinergic signaling in effort-based decision making such that blockade of muscarinic cholinergic receptors within the NAc induced effort averseness in rats (Nunes et al.
Finally, it has recently been demonstrated that delta-opioid receptors on cholinergic interneurons within the NAc shell modulate the activity of those interneurons. This influences dopaminergic D1 receptor-expressing projection neurons (medium spiny neurons) and affects stimulus-guided choice between available actions (Laurent et al. 2014). Nicotinic receptors are also involved in cost processing during decision making. For example, a recent study (Hosking et al. 2014) showed that nicotine induces shifts in rat decision making from high-reward/high-cost choices to low-reward/low-cost choices when the costs are in the form of either cognitive effort or time delay in reward delivery. However, it must be noted that the role of nicotinic receptors is region dependent, as a recent work (Mendez et al. 2013) revealed that in some regions nicotinic receptor binding is linearly related to the choice of large delayed reward whereas in some other regions the relationship is inverse.

Finally, the orexin system also appears to have a role in some aspects of decision making. Orexinergic neurons are a small population of neurons originating from lateral hypothalamus and adjacent areas innervating large but disperse areas of the brain including cortical and subcortical areas involved in decision making (Peyron et al. 1998). The orexin system is increasingly emerging as an important player in the regulation of reward seeking and motivation (Aston-Jones et al. 2009; Harris et al. 2005; Hollander et al. 2008; Mahler et al. 2014; Sakurai 2014). Although given the role of orexin in regulating reward-seeking behavior and motivation one might expect a role in decision processes as well, work investigating a possible involvement of the orexin system in regulating decision making has been scarce and limited to narcoleptic patients. These studies have tested different kinds of decision tasks in narcoleptic patients and have shown that these patients have impairments in decision making under ambiguity and are more risk prone in risk-based decision tasks (Bayard et al. 2011, 2013; Bayard and Dauvilliers 2013; Delazer et al. 2011).

**Foraging, Exploration, and Novelty Seeking**

The majority of laboratory experiments investigating decision making use decision tasks in which the subjects have to choose one of a few available options that are usually restricted to two. However, in our natural environment, the decisions we make are not always among a few well-defined options. There are situations with uncertainty about the number of options and their potential values, costs, and risks along with some other options known from previous experiences of our own or others. For decades, there have been attempts particularly among ecologists and psychologists to explain these challenging but important aspects of animal behavior in real-world decision making settings (Kamil and Roitblat 1985); only recently, however, have neuroscientists started to study neural bases of foraging and exploratory decision making. The exploration behavior gives the animal the chance to explore possible alternative courses of action that might have potentially high values compared with known alternatives. Most animals including the rodents and primates have an inherent tendency to explore novel environments and objects in their reach. Interestingly, this inherent tendency to preferentially explore novel objects and locations in the presence of already familiar objects or locations has been utilized in rodent tests of recognition memory (Ennaceur and Delacour 1988). The novelty preference in this test necessitates an intact memory of a familiar object or location. This test has gained popularity since it results from the intrinsically natural behavior of the animal and does not rely on externally provided reinforcement for learning and memory. Novelty-seeking behavior has been used to address memory (De Filippis et al. 2014; Khani and Rainer 2012) and motivation (Adriani et al. 2012) and is connected to risk taking (Laviola et al. 2003; Ruocco et al. 2014; Wang et al. 2015) and impulsivity (Koot et al. 2009). Novelty exploration poses a significant challenge in an uncertain environment between conflicting strategies of exploiting available information and gathering potentially more valuable new information. This exploratory behavior is important for strategic decision making to efficiently profit from food and other resources that might be found in unknown environments. An efficient decision maker will solve the so-called “exploitation-exploration” dilemma by exerting a balanced trade-off between choosing options known from past experiences and exploring unknown options. It is suggested that humans use at least two different strategies to solve this dilemma: a directed strategy in which choices are explicitly biased toward information seeking and a random strategy in which decision noise leads to exploration by chance (Wilson et al. 2014). Wilson et al. (2014) used a horizon task in which participants were allowed to make a single choice (horizon 1) or six sequential choices (horizon 6). They showed that subjects were more information seeking and had higher decision noise with the longer horizon in which they were allowed to make a number of sequential choices, supporting the hypothesis that humans use both strategies to solve the exploration-exploitation dilemma and suggesting that both information seeking and choice variability can be controlled and used in the service of exploration.

It has been demonstrated that dopaminergic genes predict individual differences in exploration vs. exploitation such that two genes controlling striatal dopamine function, DARPP-32 (also called PPPIR1B) and DRD2, are associated with exploitative learning whereas a gene primarily controlling prefrontal DA function (COMT) is associated with exploratory decisions that are made proportional to Bayesian uncertainty about the possibility that other choices might produce better outcome than those of the status quo (Frank et al. 2009). Neuroimaging studies in human subjects performing a gambling task revealed that the FPC and intraparietal sulcus are preferentially active during exploratory decisions while the striatum and vmPFC are involved in value-based exploitative decision making (Daw et al. 2006). As discussed above, the FPC is also involved in the encoding of unchosen options, the so-called “counterfactual choice” (Boorman et al. 2011; Rushworth et al. 2011). This is consistent with its role during exploratory decision making and suggests the neural activity related to exploring other options even after the choice. This is important, as reevaluating current choices is often necessary in the face of changing environments or changing needs. Another study has distinguished distinct brain regions involved in exploratory decision making based on information type. Wang and Voss (2014) used distinct information types available during contextual association learning along with manipulations of exploratory decision making to identify neural activity associated with information-based decisions. They identified two distinct neural correlates: a hippocampal-prefrontal contribution to advantageous deci-
sions based on immediately available novel information and a dorsal striatal contribution to advantageous decisions based on the sum total available information. In monkeys, single-unit recordings have implicated posterior cingulate cortex (PCC) in coding relative value of exploring novel resources compared with exploiting current resource (Hayden et al. 2011; Pearson et al. 2009). Activity in the supplementary eye field has also been implicated in encouraging monkeys to explore alternative decision making strategies (Donahue et al. 2013). Within basal ganglia, it has been suggested that a subset of neurons in globus pallidus internus, the primary output nucleus of the basal ganglia, are initially permissive in new settings to allow exploration of a variety of choices such that decreased firing rate predicts exploratory behavior. These neurons increase their firing rate once a profitable choice is identified and encourage exploitation of that profitable choice (Sheth et al. 2011). Foraging behavior and exploratory decision making by nature bear many degrees of freedom, resulting in differential conclusions depending on the experimental design and disagreements on key brain regions involved in processing decision making in ecological settings. In addition to the brain areas discussed, some other work has revealed neural correlates of foraging in areas different from previous ones. Rushworth and colleagues have used a sequential decision making task to compare neural correlates of comparative decision making with those of foraging. They showed that the vmPFC encodes the value of well-defined options whereas the ACC encodes the average value of the foraging environment and the cost of foraging (Kolling et al. 2012). In another experiment, they showed that activity in the vmPFC and dorsal ACC (dACC) was tied to the current choice and the best long-term option, respectively, suggesting that the vmPFC and dACC adopt choice and default (invariant) reference frames during sequential multialternative choice (Boorman et al. 2013). Consistent with this finding, it has been shown that lesion of the ACC attenuates foraging behavior in rodents (Li et al. 2012). This account of the neural mechanisms of foraging in which the dACC plays a central role has been challenged. In a recent study, Shenhav et al. (2014) showed that the dACC is associated with foraging value only when foraging value is confounded with choice difficulty and the dissociation of two variables revealed that the dACC recruitment is only accounted by choice difficulty but not the value of foraging.

![Diagram of Decision Making](image_url)

**Fig. 4.** Summary of major neural and neurochemical elements governing decision making processing. In this representation, the execution of choice is separated from the evaluation since the elements in this part are often involved in motor aspects of decision making. It is evaluation of choices that guides decision making, which itself is influenced by several factors. This representation is meant to show key players involved in decision making rather than listing an exhaustive number of studies implicating various neural and neurochemical elements of the brain in processing different aspects of decision making. ACC, anterior cingulate cortex; ACh, acetylcholine; BLA, basolateral amygdala; DA, dopamine; dlPFC, dorsolateral prefrontal cortex; DS, dorsal striatum; FPC, frontopolar cortex; IOFC, lateral orbitofrontal cortex; mOFC, medial orbitofrontal cortex; NA, norepinephrine; NAc, nucleus accumbens; OFC, orbitofrontal cortex; PCC, posterior parietal cortex; PFC, prefrontal cortex; SMA, supplementary motor area; SN, substantia nigra; STN, subthalamic nucleus; vmPFC, ventromedial prefrontal cortex; VP, ventral pallidum.
Concluding Remarks

Decision making is an adaptive behavior that ranges from everyday simple choices to important decisions with long-lasting consequences. Researchers employ multiple approaches to address questions pertaining decision making. In cognitive neuroscience, these studies fall into two broad categories of perceptual decision making and reinforcement-guided decision making. The focus of the present update is the latter. However, it must be noted that these two forms of decision making share several overlapping elements. The models of reinforcement-guided decision making are based on economic (Von Neumann and Morgenstern 1944) and reinforcement learning (Sutton and Barto 1998) theories. To make a choice, several parameters including values, costs, and degrees of certainty have to be integrated to result in a decision. These decision-related parameters are processed and integrated in a large network that consists of distributed neural circuits and uses several different neurotransmitter systems to communicate. We have provided a broad review of well-established as well as emerging accounts of the neural and neurochemical basis of reinforcement-guided decision making. A summary of differential contributions of major neural and neurochemical elements governing decision making processing is depicted in Fig. 4. The process of decision making can be divided into two phases: evaluation and execution. While the neural circuitry responsible for these two phases is still emerging, a summary can be outlined as follows. In the evaluation phase, several parameters of alternative choices are integrated and compared to ensure a best course of action in any given circumstance. This necessitates information about several aspects of the choice that is often lacking. Mounting evidence converges in portraying the ACC and OFC/vmPFC at the center of brain areas carrying out the evaluation phase. This processing implicates a larger network that includes several subcortical regions. Amygdala-OFC and NAc-OFC connections provide information regarding the emotional content and the value of choice, track a recent history of reward, and facilitate adjusting choice in response to a changing environment. The value of delayed and uncertain rewards is also dependent on these circuitries and is modulated by several neuromodulators, particularly the serotonin system. The ACC (particularly dACC) receives value and cost information from cortical and subcortical regions and integrates value and cost information. In particular, the NAc provides information concerning RPEs and costs associated with a given choice. The motivation to achieve a high reward despite a large cost associated with it is mediated by DA neurotransmission in the NAc (mesolimbic pathway) and the ACC (mesocortical pathway). The value and cost of alternative or unchosen choices are also important in guiding future decision making behavior. aPFC is at the center of such valuations, although dACC and PCC are also possibly involved in this process. The information resulting from the evaluation phase becomes accessible to the brain regions involved in planning and executing the chosen action. This process is also implemented at both subcortical and cortical levels through cortico-subcortical circuitry encompassing medial and dorsolateral PFC and basal ganglia structures including dorsal striatum and the STN. Dorsal striatum is involved in action representation and action initiation, whereas dPFC uses information from the evaluation phase to plan subsequent actions and coordinate allocation of neural resources such as attention to a particular course of action. mPFC-STN communications set decision thresholds and inhibit action before the evidence accumulation passes the set threshold.

Acknowledgments

The authors thank E. Nouri for her help in preparing the figures.

Funding

This work was supported by the Swiss National Science Foundation, EURYI Grant PE0033-117106, and the University of Fribourg.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the author(s).

Author Contributions

A.K. and G.R. edited and revised manuscript; A.K. and G.R. approved final version of manuscript; A.K. prepared figures; A.K. drafted manuscript.

References


Asher DE, Craig AB, Zaldívar A, Brewer AA, Krishmar JL. A dynamic, embodied paradigm to investigate the role of serotonin in decision-making. Front Integr Neurosci 7: 78, 2013.


Ghods-Sharifi S, Floresco SB. Differential effects on effort discounting by 10.220.33.4 on October 20, 2017 http://jn.physiology.org/ Downloaded from


Metereau E, Dreher JC. The medial orbitofrontal cortex encodes a generalized unsignaled value signal during anticipation of both appetitive and aversive events. Cortex 63C: 42–54, 2014.


Neurons in the orbitofrontal cortex encode rewards and decision making.


Rudebeck PH, Bannerman DM, Rushworth MF. The distinction of different subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. Cogn Affect Behav Neurosci 8: 485–497, 2008a.


