

Contributions of the striatum to learning, motivation, and performance: an associative account

Mimi Liljeholm and John P. O'Doherty

Division of the Humanities and Social Sciences, and Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA 91125, USA

It has long been recognized that the striatum is composed of distinct functional sub-units that are part of multiple cortico-striatal-thalamic circuits. Contemporary research has focused on the contribution of striatal sub-regions to three main phenomena: learning of associations between stimuli, actions and rewards; selection between competing response alternatives; and motivational modulation of motor behavior. Recent proposals have argued for a functional division of the striatum along these lines, attributing, for example, learning to one region and performance to another. Here, we consider empirical data from human and animal studies, as well as theoretical notions from both the psychological and computational literatures, and conclude that striatal sub-regions instead differ most clearly in terms of the associations being encoded in each region.

Anatomical and functional delineations of the striatum

Early anatomical studies delineated striatal sub-regions in terms of their afferent and efferent cortical projections (Figure 1), demonstrating that the dorsolateral region of the striatum (i.e., putamen) is primarily connected to sensory and motor cortices. In contrast, a dorsomedial region (i.e., caudate) is connected with frontal and parietal association cortices, whereas the ventral striatum is connected with limbic structures, including the amygdala, hippocampus, and medial orbitofrontal and anterior cingulate cortices [1,2]. Over the past few decades, these striatal divisions have played central roles in theoretical and empirical work across psychological domains.

First, theories of associative learning, which address how relationships between stimuli, actions, and rewards become encoded in the brain, have attributed different types of associative learning to distinct dorsal and ventral regions of the striatum [3,4]. Dissociable dorsal regions have also been identified by research that contrasts automatic performance of well-learned motor-programs with tasks that require high-level 'executive' attention or cognitive control [5,6]. In particular, in the motor-skill literature, medial and lateral regions of the dorsal striatum are often reported to be involved in early learning and well-trained performance, respectively [7–10]. More recently, learning versus performance of motor behavior has instead

been attributed to ventral versus dorsal striatal regions; specifically, it has been proposed that, whereas the ventral striatum supports both learning and performance, the dorsal striatum is only critical for performance [11]. Others have postulated a dorsal-ventral distinction with respect to how incentives modulate performance, arguing that the ventral striatum encodes motivational variables and communicates their significance to dorsal regions responsible for response implementation [12,13]. In the present review, we discuss key findings from this broad and divergent literature and contrast accounts that delineate striatal sub-regions in terms of learning, performance, or motivation with theories that emphasize the content and nature of associative encoding.

Learning and the striatum

An extensive body of work has focused on the role of the striatum in facilitating two different types of associative learning: Pavlovian learning, in which, through repeated pairings, initially neutral conditioned stimuli (CSs) come to elicit reflexive behaviors in anticipation of the subsequent occurrence of appetitive or aversive events, and instrumental learning in which an organism learns to perform actions that increase the probability of obtaining reward or avoiding punishers [14]. Instrumental learning is further divided into goal-directed learning, which is driven by representations of the outcomes of actions – their value and causal antecedents, and habit learning, through which actions come to be automatically elicited by the stimulus environment, without any explicit reference to their consequences [15].

Considerable evidence has amassed to implicate the ventral striatum (VS) in Pavlovian learning: transient dopamine (DA) release in the VS in response to primary food rewards shifts, across training, to the onset of reward-predictive cues, and CSs that signal food reward produce changes in neuronal firing patterns in the VS [16,17]. In contrast, different sub-regions of the dorsal striatum appear to be involved in habitual and goal-directed instrumental conditioning, respectively. In rodents, lesions of the lateral dorsal striatum (DLS) disrupt acquisition of habits, whereas lesions to the medial part of the dorsal striatum (DMS) impair goal-directed learning [18–20]. Likewise, in humans, activity in the DMS has been found to be correlated with computations of action-outcome contingency, a

Corresponding authors: Liljeholm, M. (mlil@hss.caltech.edu); O'Doherty, J.P. (jdoherty@caltech.edu)

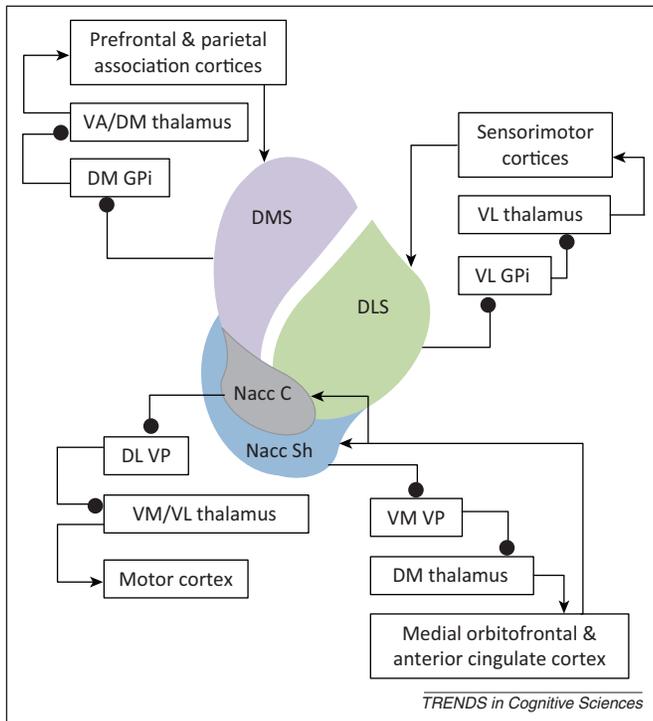


Figure 1. Schematic representation of corticostriatal connections based on [1,2,87,88]. Different cortical areas project to different sub-regions of the striatum, which then project back to respective cortical areas via the internal segment of the globus pallidus (GPI) and the thalamus (direct pathway). Not shown projections include those from different striatal sub-regions to distinct areas in the substantia nigra pars reticulata (SNr) and the external segment of the globus pallidus, as well as those from distinct midbrain nuclei (e.g., substantia nigra compacta and ventral tegmental area) to different striatal sub-regions. Moreover, in the ventral striatum, the Nacc shell, but not the core, projects heavily to the amygdala and lateral hypothalamus, whereas both the shell and core receive inputs from limbic regions (e.g., amygdala and hippocampus). Circle, inhibitory connection; Arrow, excitatory connection; DMS, dorsomedial striatum; DLS, dorsolateral striatum; GPI, internal segment of globus pallidus; VP, ventral pallidum; VA, ventral anterior; DM, dorsomedial; VL, ventrolateral; VM, ventromedial; Nacc C, nucleus accumbens core; Nacc Sh, nucleus accumbens shell.

hallmark of goal-directed learning, whereas activity in a region of right posterior DLS was found to track the behavioral development of habits (Figure 2A) [21–23].

Computational approaches to understanding the functions of the striatum are dominated by reinforcement-learning (RL) theory [24]. In one class of RL algorithms called ‘model-free’ (referring to the absence of an internal model of the world), a reward prediction error (RPE) signal is used to incrementally update reward expectations assigned to particular states of the world or to actions available in those states [25]. One RL model initially proposed as an account of striatal function is the actor/critic model [26], in which a critic module learns to anticipate rewards associated with various states of the world, analogous to Pavlovian conditioned expectations, whereas an actor module learns a policy corresponding to the probability of performing a particular action given some state, analogous to learning instrumental actions. Importantly, in this model, the RPE signals generated by the critic are used to update both the state-based reward expectations in the critic and the action probabilities in the actor. In support of this view, human fMRI studies have found that VS activity correlates with RPEs during tasks that feature exclusively Pavlovian reward associations

[27,28], consistent with a role for this region in implementing the critic, whereas tasks involving instrumental actions have been shown to recruit both ventral and dorsal striatum [27,29,30].

A major limitation of the actor/critic model is that it cannot account for the known differences between goal-directed and habitual instrumental actions, and the differential functions of the DMS and DLS in supporting these mechanisms. Specifically, the actor-critic model, using a general appetitive RPE signal, is entirely model-free, failing to provide an account of goal-directed performance and its implementation by the DMS. This shortcoming has been addressed by the proposal that goal-directed instrumental behavior can be accounted for by means of a ‘model-based’ type of RL, in which the agent encodes a rich model of the transition structure between states of the world, and uses this model, alongside knowledge of the current value of available outcomes, to perform on-line computations of the expected future value of taking particular actions [25]. In spite of the conceptual appeal of mapping quantitative model-based and model-free RL signals to the DMS and DLS respectively, very few human studies have empirically assessed this hypothesis thus far. One such study found evidence in support of the postulated computational dissociation [31], whereas another study, using a similar design, instead found evidence for a linear mix of model-based and model-free signals within the same overlapping areas [32]. Further work is needed to ascertain the extent to which model-based and model-free RL computations adequately capture the differential contributions of DMS and DLS to goal-directed and habitual learning, respectively.

Motor performance

There is considerable evidence to implicate the ventral striatum in generating skeletomotor reflexes elicited by Pavlovian cues [33,34]. Lesions as well as transient inactivation of the VS significantly impair previously acquired conditioned responses (CRs) to food-paired CSs: In particular, a medial part of the nucleus accumbens (Nacc) called the core, distinct from a more lateral part called the shell (Figure 1), has been shown to mediate the retrieval and expression of CS-US associations [33,34].

A large body of research has also implicated the dorsal striatum in the implementation of already learned instrumental motor behaviors, often with dissociations emerging between the DLS and DMS [7–10]. For example, using a serial reaction time (SRT) task, in which participants respond to a sequence of consecutively presented stimuli, several neuroimaging studies have reported that, whereas the DMS appears to be active during learning of novel sequences, the DLS is active during performance of well-learned sequences [7,8] (but see [35] for evidence of learning-related decreases in DLS activity). Notably, neurophysiological studies in non-human primates [9], as well as in rodents [10], have also found dissociable contributions of the DMS and DLS to early versus late stages of training.

The DLS and DMS also appear to differ in their contribution to the inhibition of competing, but incorrect responses, a process that is generally thought to involve

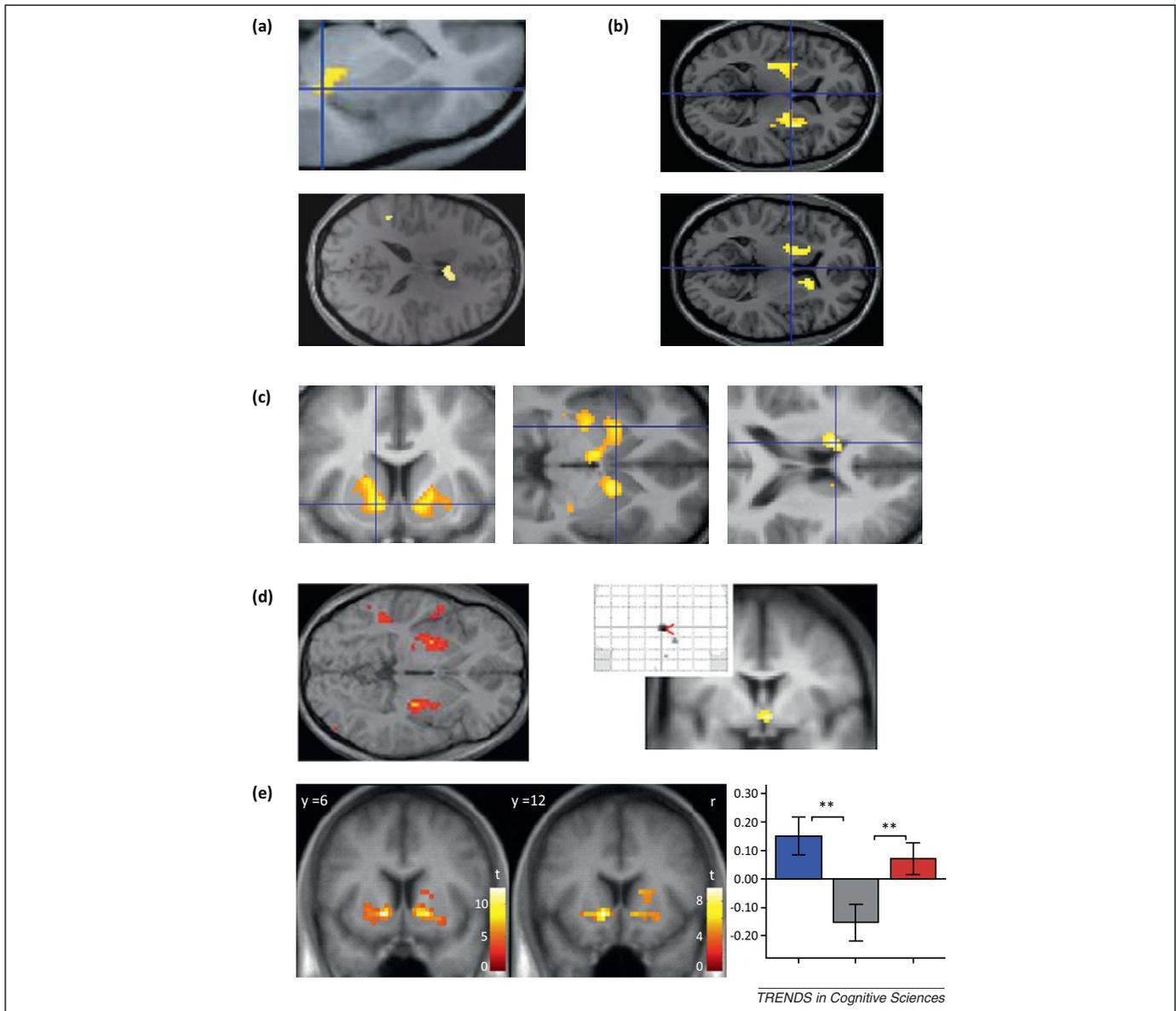


Figure 2. Human neuroimaging studies that indicate differential contributions of striatal sub-regions in associative encoding. **(a)** Striatal activity tracking the development of habits in posterior lateral striatum (top) and during goal-directed instrumental performance in anterior DMS (bottom). Reproduced, with permission, from [23] and [22], respectively. **(b)** Activity in the lateral striatum (top) as participants executed a well-trained motor sequence and in the DMS (bottom) as participants planned performance of a self-generated, novel, motor sequence. Reproduced, with permission, from [5]. No effects were found in the DMS when participants planned performance of a well-trained sequence (condition not shown here). **(c)** Effects in the VS (left) for the conjunction of high versus low incentives and correlation with performance levels. Activity in the lateral striatum is correlated with VS activity when the task entails high demands on motor performance (center), while activity in the medial striatum (right) correlates with the VS signal during high demands on cognitive performance. Reproduced, with permission, from [13]. **(d)** Imaging effects in the lateral VS for specific PIT (left) and in the medial VS for general PIT (right). Reproduced, with permission, from [47] and [46] respectively. **(e)** BOLD responses in the VS in anticipation of monetary gain (left) and loss (right), with % signal change shown in bar graph on far right, for gain (blue), no outcome (gray), and loss (red). Reproduced, with permission, from [64].

a voluntary, cognitive, suppression of automatic responding. Response inhibition is commonly studied using the Go/No Go task, in which an infrequent (No Go) stimulus signals that performance of an action that is usually rewarded will result in the omission of reward or in punishment. Neuroimaging research has implicated the DMS, more strongly than the DLS, in inhibiting responding on No Go trials [36,37]. Indeed, numerous studies have found selective involvement of the DMS in various tasks that require cognitive control and working memory [6,35,38], consistent with the strong anatomical connections of this area to pre-frontal and parietal association cortices. In Box 1, we relate the literature on skill-learning and cognitive control to that discussed in the above section on associative

learning. Additional evidence for the specialized contributions of the DLS and DMS to automatic and cognitively controlled performance, respectively, comes from investigations of neuropathology, in particular from studies on Parkinson's disease (Box 2).

One interpretation of the motor-skill literature is that the DMS and DLS can be distinguished in terms of their respective contributions to the acquisition versus performance of motor behavior [10]. However, this hypothesis is challenged by the finding that both lesions and transient inactivation of the DMS abolish the sensitivity of previously acquired actions to outcome devaluation and contingency degradation – behavioral assays of goal-directed performance [20]. Thus, DMS disruptions impair the

Box 1. The relationship between instrumental control strategies and the multiple memory systems framework

Research on skill learning and cognitive control is often guided by a 'multiple memory systems' framework that contrasts declarative memory, which provides flexible and explicit access to semantic and episodic content, but which requires conscious awareness, with memory of how to implement procedures (e.g., how to perform a sequence of actions), which is, or can become, automatic and subconscious [89]. It does not seem implausible that goal-directed deliberation of the utilities and casual antecedents of future outcomes is declarative nor that habitual and Pavlovian processes are procedural. There are, however, some important differences between the neural substrates identified by research on multiple memory systems and that addressing instrumental control strategies. In particular, declarative processes appear to depend on hippocampal areas, whereas goal-directed learning *per se* does not. It is also worth noting that there is no direct behavioral evidence supporting the equivalence of goal-directed and declarative, or of procedural and habitual, processes: a strong resistance to dual-task interference, the behavioral test used to identify automatic procedural performance [35], has not been empirically related to insensitivity to outcome devaluation and contingency degradation – defining features of habitual performance. Future work is needed to determine the exact relationship between instrumental control systems and multiple memory systems.

expression of goal-directed behavior, suggesting that this structure plays a critical role during performance. Likewise, the proposal that the dorsal striatum is critical only for performance, whereas the ventral striatum supports both learning and performance, of instrumental actions [11] is challenged by the finding that blockage of NMDA receptors in the DMS during action-outcome learning abolishes sensitivity to outcome devaluation in subsequent tests [19].

Motivation

Another function attributed to the striatum, and to the ventral striatum in particular, is that of motivation. Cues that indicate that a certain amount of reward is available given successful performance of an instrumental action, or even of a complex cognitive task, elicit increases in VS activity proportional to the amount of signaled reward and these signals correlate with the degree of performance enhancement found for larger compared to smaller rewards [12,13]. Paradoxically, whereas increasing rewards tend generally to improve performance, the opportunity to earn very large rewards has also been shown to have a deleterious influence, a phenomenon known in the psychological literature as choking. Recent neuroimaging studies have implicated the VS in these detrimental, as well as in the facilitating, effects of incentives on performance [39,40].

Cues that signal reward delivery independently of whether or not an instrumental action is performed can nevertheless invigorate instrumental performance, a phenomenon termed Pavlovian-instrumental transfer (PIT) [41,42]. These effects also appear to be largely dependent on the VS [43–45]. For example, amphetamine injection into the Nacc enhances PIT, without affecting base rates of instrumental responding [45]. Importantly, PIT effects emerge even when the instrumental action earns a different reward than that signaled by the cue and are attenuated by general motivational shifts from hunger to satiety [42], suggesting that the cue induces a general motivational state (i.e., general PIT). However, under certain training

Box 2. Striatal function and Parkinson's disease

Parkinson's disease (PD) is a neurodegenerative disorder, in which a loss of DA-producing cells in the substantia nigra (SN) impacts dorsal striatal DA function, with particularly severe DA depletion occurring throughout the putamen and in the most dorsal aspects of the caudate [90]. PD patients are impaired on a range of cognitive and sensorimotor tasks, including probabilistic classification learning and conceptual set-shifting [91], and also exhibit clear deficits in reward processing [92]. With respect to the SRT task, patients exhibit longer reaction times than healthy controls, while being relatively spared on performance accuracy as well as on declarative encoding of sequences [93]. Moreover, even when able to learn (i.e., accurately perform) a complex novel sequence, PD patients are impaired at achieving automaticity, as assessed by dual-task performance [94]. Similar results have been found using targeted lesions in rodents: dorsal (but not ventral) striatal NMDA lesions produce clear deficits in SRT performance, with impairments being more severe for reaction times than for accuracy, and more pronounced for DLS than for DMS [95]. Importantly, the reverse pattern of results was observed in a radial arm maze task, with significant impairments emerging for ventral but not dorsal striatal lesions, ruling out a general inability to initiate sequential locomotor acts as an explanation for SRT performance [95]. This finding suggests that reaction time impairments on the SRT task, due to dorsal striatal dysfunction or damage, reflect deficits in stimulus-based action selection, rather than action initiation. As with the SRT task, PD patients are impaired on Go/No Go responding, with deficits being reported for both response times [96] and accuracy [97], and with differences emerging between PD patients and healthy controls in DMS activity during Go/No Go performance [98]. Degrees of DMS dysfunction in PD patients have also been shown to correlate significantly with impairment on other measures of executive function, such as the Stroop test [99].

conditions, PIT effects exhibit a clear selectivity, such that instrumental responding is enhanced specifically for an action that earns the same reward as that signaled by the Pavlovian cue, suggesting the involvement of outcome-specific representations (i.e., specific PIT). Findings from rodent lesion and inactivation studies suggest that the Nacc shell and core may mediate specific and general PIT, respectively [41]. More recently, the involvement of the medial VS in a form of PIT that may depend on general motivational processes [46], and of the ventrolateral striatum in specific PIT [47], has been demonstrated in human neuroimaging studies (Figure 2b). A more detailed comparison of the functional anatomy of humans and rodents is provided in Box 3.

Another important function recently attributed to the ventral striatum is the hedonic evaluation of stimuli, termed 'liking', which is commonly assessed using measures of affective facial reactions [48]. Unlike PIT and a range of other reward-oriented behaviors, including approach and consumption, behavioral expressions of liking are unaffected by amphetamine injection into the Nacc [43,44]. Instead, such responses are altered by blockage or stimulation of Nacc opioid receptors [44,49], suggesting that dissociable neurobiological substrates in the VS mediate motivational and hedonic processes. Notably, although both dopaminergic and opioidergic manipulations of the Nacc modulate the firing of VP neurons in response to (reward proximal) Pavlovian cues, only opioid manipulations alter VP firing in response to unconditioned stimuli, suggesting that the separation of motivational and hedonic processes is preserved throughout the Nacc-VP circuit [44].

Box 3. Functional anatomy in humans and rodents

A remarkable degree of homology between the functional organization of human and rodent brains has been demonstrated [50]. For instance, the prelimbic cortex, identified in the rat as playing a role in the acquisition and performance of goal-directed actions [15,100], bears strong functional resemblance to the region of the ventromedial prefrontal cortex (vmPFC) implicated in goal-directed computations in human fMRI [101,102]. Furthermore, there seem to be considerable functional homologies within the striatum. In both humans and rats, the ventral striatum has been implicated in Pavlovian processes and in Pavlovian to instrumental transfer [27,33,41,43,46,47]. Moreover, within the dorsal striatum in both species, medial regions are implicated in goal-directed learning [19–22,31], whereas lateral regions are implicated in habit learning [18,23,31]. However, there may also be some differences in the precise location within the medial and lateral parts of the dorsal striatum between species. For example, whereas goal-directed performance in humans correlates with activity in an anterior part of the DMS (see bottom of Figure 2a), only disruptions of the posterior, but not the anterior, DMS abolish goal-directed performance in rodents (right panel in Figure 1a) [20]. Likewise, whereas habitual performance in rodents depends on central areas of the DLS (left panel in Figure 1b) [18], evidence from human neuroimaging studies to date have implicated a much more posterior area of the lateral putamen (top of Figure 1a) [23]. Regional differences are also apparent with respect to the contributions of ventral striatal regions to specific PIT: whereas in the rodent literature the shell of the nucleus accumbens has been found to mediate specific PIT effects (Figure 1b) [41], human neuroimaging studies have instead reported the involvement of more lateral parts of the ventral striatum outside of the nucleus accumbens proper in this process (Figure 2d) [46,47]. Of course, because the effects of lesions to lateral aspects of the rodent VS have not been assessed, the possibility remains that this area mediates specific PIT in rodents, as well as humans. Thus, although there are broad similarities across species in the corticostriatal circuits involved, in some cases more research is needed to establish precise homologies. The emergence of high-resolution fMRI as a research tool might help considerably in segregating function between different sub-regions of the human striatum at a level of specificity currently achieved only in rodent studies.

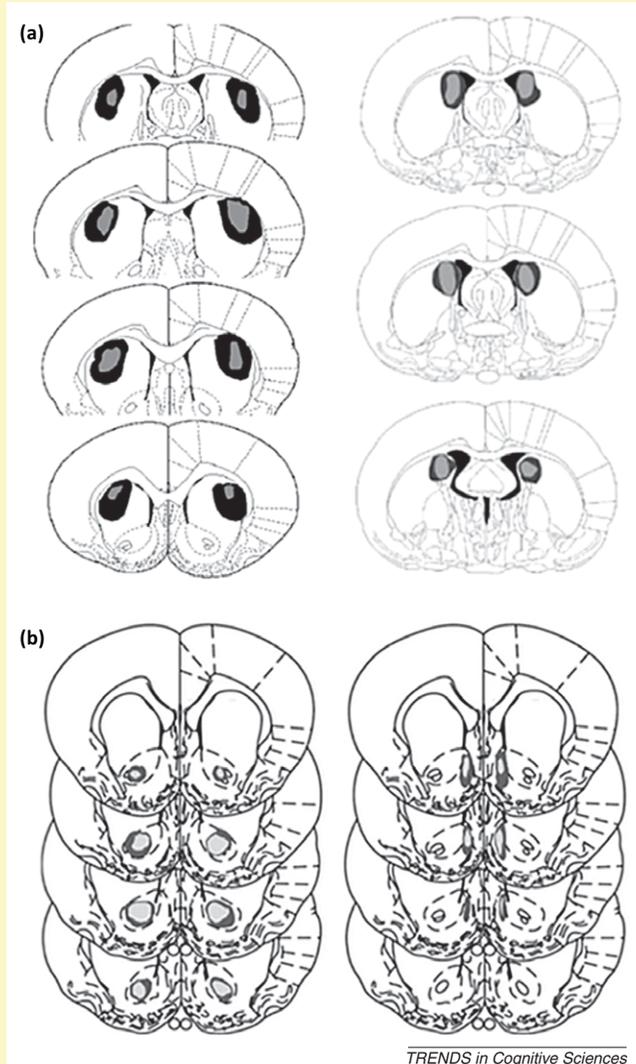


Figure 1. Schematic representations of excitotoxic striatal lesions of rodent brain. (a) Lesions of the DLS (left) that abolish habitual performance and lesions of the posterior DMS (right) that abolish goal-directed performance. Reproduced, with permission, from [18] and [20], respectively. (b) Lesions of the core (left) and shell (right) of the nucleus accumbens, respectively abolishing outcome general and outcome specific PIT. Reproduced, with permission, from [41].

An associative account of striatal function

The evidence reviewed here has implicated the ventral and dorsal striatum (both the DLS and DMS) in the learning as well as the performance of reward-related behaviors. It is unlikely therefore, that these regions differ functionally in terms of their respective contributions to learning vs performance [10,11]. Rather, a more parsimonious interpretation is that striatal regions support dissociable associative learning strategies that may respectively dominate at various stages of training, depending on the task [3,50,51]. Specifically, the ventral striatum is involved in the encoding of Pavlovian associations, supporting generation of conditioned skeletomotor responses, whereas the DMS is involved in the encoding of goal-directed instrumental actions and the DLS in the encoding of habitual stimulus-response associations. From this perspective, selective activation of the VS or DMS during early stages of

training reflects the respective dominance of Pavlovian and goal-directed instrumental processes, rather than learning *per se*.

Findings implicating the ventral striatum in incentive-based performance [12,13,39,40] can arguably also be accounted for in terms of the role of this structure in the expression of Pavlovian conditioned responses. For example, performance of an instrumental action that involves approach towards a food location may be facilitated by the presence of Pavlovian cues that elicit compatible conditioned reflexes (i.e., directed at the same location). Conversely, performance of highly skilled motor behavior or of instrumental responses that necessitate approach towards aversive stimuli might be impaired by incompatible reflexes elicited by Pavlovian cues [39]. Another potential means by which Pavlovian associations might produce both facilitatory and detrimental incentive effects on performance is

through the elicitation of habits. Specifically, Pavlovian retrieval of sensory-specific features of unconditioned stimuli might evoke stimulus representations that have been previously linked to particular instrumental responses through stimulus-response learning and that, consequently, elicit habitual performance of those responses at the point of Pavlovian retrieval [52]. Depending on whether such responses are compatible or incompatible with the instrumental actions needed to obtain the reward, a behavioral effect of either facilitation or impairment might occur.

Finally, Pavlovian retrieval of affective aspects of unconditioned stimuli contributes to the elicitation of hedonic, emotional, conditioned responses indicative of 'liking' [48]. Indeed, in this capacity, Pavlovian processes may also play a role in the estimation of outcome utility, central to accounts of goal-directed instrumental performance. This notion is particularly compelling given that CRs themselves exhibit sensitivity to outcome devaluation procedures, as we discuss further in the section below. It is also consistent with the strong projections between the VS and the medial orbitofrontal cortex (mOFC), an area well known for its involvement in utility estimation [53,54].

Challenges and further directions

RL theories of behavioral control attempt to characterize the instantiation of, and arbitration between, various associative processes and, further, to map such processes – in the form of distinct algorithms – to different striatal subregions. Although there is mounting evidence in favor of this approach, a number of key challenges still remain.

First among these is the question whether Pavlovian signals in the ventral striatum are model-free, model-based, or both. Current computational accounts of Pavlovian learning in the ventral striatum propose that such learning is model-free: that is, based on general appetitive RPE signals that are void of specific outcome representations and, thus, insensitive to changes in outcome value. This notion is greatly challenged by the fact that Pavlovian CRs, as well as BOLD signals in the VS, show clear sensitivity to outcome-specific devaluation [55–57]. Attempts to resolve this apparent inconsistency have included the proposal that preparatory (e.g., approach) and consummatory (e.g., chewing) CRs may be model-free and model-based, respectively, and that these different algorithms may be implemented by the core and shell of the Nacc, respectively [58]. Although promising, this revised RL account faces some problems; most notably, the Nacc core and shell have both been shown to be necessary for the effects of outcome-specific devaluation on preparatory CRs [56,57]. Nevertheless, it is clear that humans, as well as other animals, are capable of learning about the specific features of Pavlovian outcomes and that the VS appears to play a role in such effects.

A second question concerns the role of the striatum in aversive learning and in processing novel stimuli. Developing an understanding of the role of the striatum in aversive learning represents a major challenge. RL theory has focused almost exclusively on the role of reward in Pavlovian and instrumental processes. Indeed, because of our focus on such computational accounts, our own discussion has been geared towards appetitive learning – a bias

that is also explained, in part, by a general emphasis in the literature on reward processing in the striatum, with processing of aversive events being primarily attributed to other regions, such as the amygdala, anterior insula and lateral OFC [59–61]. However, the neuroimaging literature is profoundly inconsistent on this point, with some studies reporting increased VS activity in aversive contexts (Figure 2c) [62–64] and others reporting decreasing activity in this area during the prediction, learning, and receipt of aversive outcomes [65,66]. Likewise, whereas some studies have reported that aversive stimuli inhibit the DA activity of midbrain neurons (e.g., [67]), others have found that they elicit phasic DA release in the VS (e.g., [68]).

One possible reason for these variable findings might be that ventral striatal responses are strongly contextually dependent. A clear example of context dependent value encoding comes from a study in which the firing of ventral pallidal (VP) neurons in response to an intense salt solution was measured in rodents while in a normal homeostatic state versus a salt-deprived state. Behavioral measures of hedonic processing revealed that the solution was strongly aversive when rats were in a normal state, but became pleasant in the salt-deprived state. Intriguingly, the response patterns of VP neurons closely tracked such behavioral changes, showing a dramatic increase in response to the salt-solution in the deprived relative to the normal state [69]. Thus, the same stimulus was perceived, and neurally encoded, as both pleasant and aversive depending on the subject's internal context. Precisely how such context-dependent encoding effects become manifest within the striatum is going to be an important area of future research.

In addition to aversive and appetitive encoding, DA neurons across the mesolimbic, mesocortical and nigrostriatal pathways have been shown to respond phasically to novel environmental stimuli [70], regardless of their particular valence (i.e., appetitive, aversive, or neutral). In the VS specifically, responses to novel stimuli have been shown with fast-scan voltammetry and other techniques measuring extra-cellular DA concentrations, as well as with single unit recordings and fMRI [71–73]. An important aspect of encoding novel events is that they may serve as a basis for exploration. In this sense, it behooves the organism to effectively treat novelty as a rewarding event, thus promoting approach towards and search of unfamiliar, but potentially richly rewarding, environments. Indeed, some behavioral evidence from rodents suggests that novelty may serve as an instrumental reinforcer, such that rats will press a lever that produces an apparently neutral light stimulus more than a lever that does not yield any outcome [74]. Several modified RL algorithms have been proposed that incorporate novel event signaling, either as a surrogate of reward or as a component of the estimated state value [75].

Another area where outstanding questions remain concerns corticostriatal interactions. Although there is overwhelming evidence for a role of the DLS in performance of well-learned motor programs [5,7–10,18], consistent with the characterization of this area by RL theory as the site where habits are ultimately stored and expressed, some

data indicate that over-trained responses can be independent of the DLS specifically [76] and of DA more generally [77]. On these grounds, it has been suggested that reinforcement-learning in the striatum provides a basis for successful Hebbian learning in sensory and premotor cortices and that, with extended training, control is transferred to these less plastic, but considerably faster, cortico-cortical projections [78]. Additional support for this view comes from neuroimaging studies showing that, with extremely extended training (i.e., several weeks), slowly evolving BOLD signals in the primary motor cortex (M1) begin to discriminate between practiced and novel sequences [79].

Conversely, tasks such as deductive reasoning and problem solving, which are known to depend largely on high-level association cortices and which have no obvious connection to reward learning, seem nonetheless to recruit strongly the DMS [80,81], suggesting that this structure implements far more complex functions than those outlined by RL theory. Generally, these issues highlight the importance of considering the interplay between the striatum and cortex in accounting for the specialization of striatal sub-regions.

Another important consideration is whether striatal sub-regions differ in terms of the mechanism underlying selection between alternative responses. In RL theory, one simple way to implement action selection in either a model-based or model-free learner is to use a soft-max distribution [24,25], in which a free parameter controls the degree to which choices are biased towards the highest valued action. However, in many cases, the basis for exploration of non-optimal response alternatives, permitting discovery of actions that are more rewarding than those sampled thus far, is likely more principled than that afforded by the soft-max rule. For example, exploratory sampling might be guided by uncertainty about the relationships between actions and rewards [82]. One possibility is that model-based processes implement selection based on such relative uncertainty estimation, whereas the habit system uses the blunter soft-max rule. Alternatively, the selection mechanism for habitual, as well as Pavlovian, systems might be better characterized by simple drift diffusion models (DDM) [83], in which, at every instance, noisy 'evidence' is accumulated for each response alternative until a threshold, serving as the decision criterion, is reached. DDMs have been shown to successfully capture perceptual [84] and value-based [85] decision-making, as well as the firing rates of neurons in the lateral intraparietal area of the monkey brain [84]. A major avenue for future work will be to determine how striatal regions differ, or are similar, in their implementation of response selection, as well as to develop a better understanding of the role of corticostriatal interactions in such response selection functions.

Concluding remarks

In this article, we have reviewed evidence implicating the striatum as a whole in a number of distinct processes underlying reinforcement-related motor behavior: in learning of both instrumental actions and Pavlovian conditioned responses, in the expression of such learned

behaviors, and in controlling the motivation to respond. We have noted that, rather than being divided along lines of learning versus performance, striatal subregions appear to implement distinct forms of associative encoding. Specifically, the ventral striatum is more involved in Pavlovian conditioned responses, whereas the dorsal striatum is involved in instrumental action. Moreover, there is a dissociation within the dorsal striatum – between medial and lateral structures – in the implementation of goal-directed and habitual instrumental strategies. Finally, through its role in the learning and expression of Pavlovian conditioned responses, rather than, perhaps, through its role in motivation *per se*, the ventral striatum supports a range of modulatory influences on instrumental performance, including general invigoration (e.g., general PIT), response selection (specific PIT), and potentially even goal-directed outcome evaluation.

The question of how dissociable striatal modules, supporting distinct associative processes, compete and cooperate is at the center of the associative account of striatal function [25,86]. Although much is now known about how striatal regions differ, much less is understood about the mechanisms by which they interact with each other and with the cortex. Future work will need to move beyond the functional segregation perspective and focus instead on characterizing how distinct circuits integrate to produce coordinated cognitive and motor behavior.

References

- 1 Lynd-Balta, E. and Haber, S.N. (1994) The organization of midbrain projections to the striatum in the primate: sensorimotor-related striatum versus ventral striatum. *Neuroscience* 59, 625–640
- 2 Zahm, D.S. *et al.* (1996) Ventral striatopallidothalamic projection: IV. Relative involvements of neurochemically distinct subterritories in the ventral pallidum and adjacent parts of the rostroventral forebrain. *J. Comp. Neurol.* 364, 340–362
- 3 Balleine, B. *et al.* (2008) Multiple forms of value learning and the function of dopamine. In *Neuroeconomics: Decision Making and the Brain* (Glimcher, P.W. *et al.*, eds), Academic Press
- 4 Balleine, B.W. *et al.* (2009) The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* 199, 43–52
- 5 Jankowski, J. *et al.* (2009) Distinct striatal regions for planning and executing novel and automated movement sequences. *Neuroimage* 44, 1369–1379
- 6 Mestres-Misse, A. *et al.* (2012) An anterior-posterior gradient of cognitive control within the dorsomedial striatum. *Neuroimage* 62, 41–47
- 7 Grafton, S.T. *et al.* (1995) Functional mapping of sequence learning in normal humans. *Cogn. Neurosci.* 7, 497–510
- 8 Jueptner, M. *et al.* (1997) Anatomy of motor learning. I. Frontal cortex and attention to action. *J. Neurophysiol.* 77, 1313–1324
- 9 Miyachi, S. *et al.* (2002) Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp. Brain Res.* 146, 122–126
- 10 Yin, H.H. *et al.* (2009) Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat. Neurosci.* 12, 333–341
- 11 Atallah, H.E. *et al.* (2007) Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.* 10, 126–131
- 12 Pessiglione, M. *et al.* (2007) How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science* 316, 904–906
- 13 Schmidt, L. *et al.* (2012) Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biol.* 10, e1001266
- 14 Hall, G. (2002) Associative structures in Pavlovian and instrumental conditioning. In *Steven's Handbook of Experimental Psychology: Learning, Motivation and Emotion* (Pashler, H. and Gallistel, R., eds), pp. 1–45, John Wiley & Sons, Inc

- 15 Balleine, B.W. and Dickinson, A. (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419
- 16 Day, J.J. et al. (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028
- 17 Day, J.J. et al. (2006) Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. *Eur. J. Neurosci.* 23, 1341–1351
- 18 Yin, H.H. et al. (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189
- 19 Yin, H.H. et al. (2005) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22, 505–512
- 20 Yin, H.H. et al. (2005) The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523
- 21 Liljeholm, M. et al. (2011) Neural correlates of instrumental contingency learning: differential effects of action-reward conjunction and disjunction. *J. Neurosci.* 31, 2474–2480
- 22 Tanaka, S.C. et al. (2008) Calculating consequences: brain systems that encode the causal effects of actions. *J. Neurosci.* 28, 6750–6755
- 23 Tricomi, E. et al. (2009) A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232
- 24 Sutton, R. and Barto, A. (1998) *Reinforcement Learning: An Introduction*, MIT Press
- 25 Daw, N.D. et al. (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711
- 26 Barto, A.C. (1995) Adaptive critics and the basal ganglia. In *Models of Information Processing in the Basal Ganglia* (Houk, J.C. et al., eds), pp. 215–232, MIT Press
- 27 O'Doherty, J. et al. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454
- 28 O'Doherty, J.P. et al. (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337
- 29 Cooper, J.C. et al. (2012) Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. *J. Cogn. Neurosci.* 24, 106–118
- 30 Glascher, J. et al. (2009) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb. Cortex* 19, 483–495
- 31 Wunderlich, K. et al. (2012) Mapping value based planning and extensively trained choice in the human brain. *Nat. Neurosci.* 15, 786–791
- 32 Daw, N.D. et al. (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215
- 33 Blaiss, C.A. and Janak, P.H. (2009) The nucleus accumbens core and shell are critical for the expression, but not the consolidation, of Pavlovian conditioned approach. *Behav. Brain Res.* 200, 22–32
- 34 Parkinson, J.A. et al. (1999) Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J. Neurosci.* 19, 2401–2411
- 35 Poldrack, R.A. et al. (2005) The neural correlates of motor skill automaticity. *J. Neurosci.* 25, 5356–5364
- 36 Chevrier, A.D. et al. (2007) Dissociation of response inhibition and performance monitoring in the stop signal task using event-related fMRI. *Hum. Brain Mapp.* 28, 1347–1358
- 37 Wager, T.D. et al. (2005) Common and unique components of response inhibition revealed by fMRI. *Neuroimage* 27, 323–340
- 38 Levy, R. et al. (1997) Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *J. Neurosci.* 17, 3870–3882
- 39 Chib, V.S. et al. (2012) Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron* 74, 582–594
- 40 Mobbs, D. et al. (2009) Choking on the money: reward-based performance decrements are associated with midbrain activity. *Psychol. Sci.* 20, 955–962
- 41 Corbit, L.H. and Balleine, B.W. (2011) The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J. Neurosci.* 31, 11786–11794
- 42 Corbit, L.H. et al. (2007) General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *Eur. J. Neurosci.* 26, 3141–3149
- 43 Berridge, K.C. et al. (2009) Dissecting components of reward: 'liking', 'wanting', and learning. *Curr. Opin. Pharmacol.* 9, 65–73
- 44 Smith, K.S. et al. (2011) Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. *Proc. Natl. Acad. Sci. U.S.A.* 108, E255–E264
- 45 Wyvell, C.L. and Berridge, K.C. (2001) Incentive sensitization by previous amphetamine exposure: increased cue-triggered 'wanting' for sucrose reward. *J. Neurosci.* 21, 7831–7840
- 46 Talmi, D. et al. (2008) Human pavlovian-instrumental transfer. *J. Neurosci.* 28, 360–368
- 47 Bray, S. et al. (2008) The neural mechanisms underlying the influence of Pavlovian cues on human decision making. *J. Neurosci.* 28, 5861–5866
- 48 Berridge, K.C. and Robinson, T.E. (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.* 28, 309–369
- 49 Wassum, K.M. et al. (2009) Distinct opioid circuits determine the palatability and the desirability of rewarding events. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12512–12517
- 50 Balleine, B.W. and O'Doherty, J.P. (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69
- 51 Yin, H.H. et al. (2008) Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* 28, 1437–1448
- 52 Balleine, B.W. and Ostlund, S.B. (2007) Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann. N. Y. Acad. Sci.* 1104, 147–171
- 53 Hare, T.A. et al. (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630
- 54 Plassmann, H. et al. (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J. Neurosci.* 30, 10799–10808
- 55 Gottfried, J.A. et al. (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107
- 56 Lex, B. and Hauber, W. (2010) The role of nucleus accumbens dopamine in outcome encoding in instrumental and Pavlovian conditioning. *Neurobiol. Learn. Mem.* 93, 283–290
- 57 Singh, T. et al. (2010) Nucleus accumbens core and shell are necessary for reinforcer devaluation effects on pavlovian conditioned responding. *Front. Integr. Neurosci.* 4, 126
- 58 Bornstein, A.M. and Daw, N.D. (2011) Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Curr. Opin. Neurobiol.* 21, 374–380
- 59 Fanselow, M.S. (2010) From contextual fear to a dynamic view of memory systems. *Trends Cogn. Sci.* 14, 7–15
- 60 Gottfried, J.A. et al. (2002) Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 10829–10837
- 61 O'Doherty, J. et al. (2001) Representation of pleasant and aversive taste in the human brain. *J. Neurophysiol.* 85, 1315–1321
- 62 Delgado, M.R. et al. (2011) Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. *Front. Neurosci.* 5, 71
- 63 Jensen, J. et al. (2003) Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257
- 64 Wrase, J. et al. (2007) Different neural systems adjust motor behavior in response to reward and punishment. *Neuroimage* 36, 1253–1262
- 65 Cooper, J.C. et al. (2009) Available alternative incentives modulate anticipatory nucleus accumbens activation. *Soc. Cogn. Affect. Neurosci.* 4, 409–416
- 66 Tom, S.M. et al. (2007) The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518
- 67 Mirenowicz, J. and Schultz, W. (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451
- 68 Budygin, E.A. et al. (2012) Aversive stimulus differentially triggers subsecond dopamine release in reward regions. *Neuroscience* 201, 331–337

- 69 Tindell, A.J. *et al.* (2006) Ventral pallidum firing codes hedonic reward: when a bad taste turns good. *J. Neurophysiol.* 96, 2399–2409
- 70 Horvitz, J.C. (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656
- 71 Ljungberg, T. *et al.* (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163
- 72 Rebec, G.V. (1998) Real-time assessments of dopamine function during behavior: single-unit recording, iontophoresis, and fast-scan cyclic voltammetry in awake, unrestrained rats. *Alcohol. Clin. Exp. Res.* 22, 32–40
- 73 Wittmann, B.C. *et al.* (2008) Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973
- 74 Reed, P. *et al.* (1996) Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination. *Anim. Learn. Behav.* 24, 38–45
- 75 Kakade, S. and Dayan, P. (2002) Dopamine: generalization and bonuses. *Neural Netw.* 15, 549–559
- 76 Turner, R.S. *et al.* (2005) Sequential motor behavior and the basal ganglia. In *The basal ganglia VIII (Advances in Behavioral Biology)* (Bolam, J.P. *et al.*, eds), pp. 563–574, Springer
- 77 Choi, W.Y. *et al.* (2005) Extended habit training reduces dopamine mediation of appetitive response expression. *J. Neurosci.* 25, 6729–6733
- 78 Ashby, F.G. *et al.* (2010) Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci.* 14, 208–215
- 79 Ungerleider, L.G. *et al.* (2002) Imaging brain plasticity during motor skill learning. *Neurobiol. Learn. Mem.* 78, 553–564
- 80 Rodriguez-Moreno, D. and Hirsch, J. (2009) The dynamics of deductive reasoning: an fMRI investigation. *Neuropsychologia* 47, 949–961
- 81 Stocco, A. and Anderson, J.R. (2008) Endogenous control and task representation: an fMRI study in algebraic problem-solving. *J. Cogn. Neurosci.* 20, 1300–1314
- 82 Badre, D. *et al.* (2012) Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron* 73, 595–607
- 83 Ratcliff, R. (1978) A theory of memory retrieval. *Psychol. Rev.* 85, 59–108
- 84 Churchland, A.K. *et al.* (2008) Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702
- 85 Kraglich, I. *et al.* (2010) Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298
- 86 Dayan, P. *et al.* (2006) The misbehavior of value and the discipline of the will. *Neural Netw.* 19, 1153–1160
- 87 Draganski, B. *et al.* (2008) Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. *J. Neurosci.* 28, 7143–7152
- 88 Zahm, D.S. and Brog, J.S. (1992) On the significance of subterritories in the ‘accumbens’ part of the rat ventral striatum. *Neuroscience* 50, 751–767
- 89 Poldrack, R.A. and Foerde, K. (2008) Category learning and the memory systems debate. *Neurosci. Biobehav. Rev.* 32, 197–205
- 90 Snow, B.J. *et al.* (2000) Pattern of dopaminergic loss in the striatum of humans with MPTP induced parkinsonism. *J. Neurol. Neurosurg. Psychiatry* 68, 313–316
- 91 Beatty, W.W. and Monson, N. (1990) Problem solving in Parkinson’s disease: comparison of performance on the Wisconsin and California Card Sorting Tests. *J. Geriatr. Psychiatry Neurol.* 3, 163–171
- 92 Bodi, N. *et al.* (2009) Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson’s patients. *Brain* 132, 2385–2395
- 93 Stefanova, E.D. *et al.* (2000) Visuomotor skill learning on serial reaction time task in patients with early Parkinson’s disease. *Mov. Disord.* 15, 1095–1103
- 94 Wu, T. and Hallett, M. (2005) A functional MRI study of automatic movements in patients with Parkinson’s disease. *Brain* 128, 2250–2259
- 95 Mair, R.G. *et al.* (2002) A double dissociation within striatum between serial reaction time and radial maze delayed nonmatching performance in rats. *J. Neurosci.* 22, 6756–6765
- 96 Cooper, J.A. *et al.* (1994) Slowed central processing in simple and go/no-go reaction time tasks in Parkinson’s disease. *Brain* 117 (Pt 3), 517–529
- 97 Bokura, H. *et al.* (2005) Event-related potentials for response inhibition in Parkinson’s disease. *Neuropsychologia* 43, 967–975
- 98 Baglio, F. *et al.* (2011) Functional brain changes in early Parkinson’s disease during motor response and motor inhibition. *Neurobiol. Aging* 32, 115–124
- 99 Nobili, F. *et al.* (2010) Cognitive-nigrostriatal relationships in de novo, drug-naive Parkinson’s disease patients: a [¹²³I]FP-CIT SPECT study. *Mov. Disord.* 25, 35–43
- 100 Ostlund, S.B. and Balleine, B.W. (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.* 25, 7763–7770
- 101 de Wit, S. *et al.* (2009) Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* 29, 11330–11338
- 102 Valentin, V.V. *et al.* (2007) Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026