The role of the basal ganglia in habit formation

Henry H. Yin* and Barbara J. Knowlton‡

Abstract | Many organisms, especially humans, are characterized by their capacity for intentional, goal-directed actions. However, similar behaviours often proceed automatically, as habitual responses to antecedent stimuli. How are goal-directed actions transformed into habitual responses? Recent work combining modern behavioural assays and neurobiological analysis of the basal ganglia has begun to yield insights into the neural basis of habit formation.

When you flip on a light switch, your behaviour could be a result of the desire for a state of illumination coupled with the belief that a certain movement will lead to it. Sometimes, however, you just turn on the light habitually, without anticipating the consequences — the very context of having arrived home in a dark room automatically triggers your reaching for the light switch. Although to the observer these two cases might appear to be similar, they differ in the extent to which they are controlled by outcome expectancy. When the light switch is known to be broken, the habit might still persist whereas the goal-directed action might not.

Intuitively, then, goal-directed actions are controlled by their consequences, habits by antecedent stimuli. But how can we translate such intuitive concepts into operationally defined terms and experimentally testable hypotheses? Here, we outline the basic conceptual framework that has emerged from the behavioural analysis of goal-directed actions and stimulus-driven habits, and integrate this framework with recent findings on the anatomy and physiology of the basal ganglia, a set of nuclei that have long been known to control voluntary behaviour. More specifically, we show that distinct networks involving the basal ganglia are the neural implementations of actions and habits, and that an understanding of these networks can illuminate findings from different levels of analysis, from the cellular and molecular mechanisms of synaptic plasticity to the conditions that favour habit formation and the development of compulsivity in various clinical disorders.

Basal ganglia and instrumental behaviours

The basal ganglia: anatomy and functions. The basal ganglia are a set of nuclei located in the cerebrum (FIG. 1). Unlike the cortex, which has excitatory, glutamatergic projection neurons, the basal ganglia contain inhibitory, GABA (γ-aminobutyric acid)-containing projection neurons. Of these projection neurons, the spiny variety belongs to the striatum (the input nucleus) and the aspiny variety belongs to the pallidum (the output nucleus)[1,2].

The striatal projection neurons are often quiescent owing to their intrinsic membrane properties[2], and when they are activated by strong and coherent inputs from the cortex (and, to a lesser extent, the thalamus), they tend to reduce the tonically active pallidal output. The outcome of this disinhibitory pathway, the most basic pathway in the basal ganglia, is the facilitation of the targeted motor network[3]. However, a different pathway, traditionally known as the 'indirect pathway', appears to exert inhibitory control over downstream thalamocortical and brainstem networks[4].

In discussing the role of the basal ganglia in behaviour, it is useful to think of them as a biological system that operates by classic selectionist principles, possessing a generator of diversity and mechanisms of selection and of differential amplification. The striatum receives massive projections from almost all cortical areas, and from the intralaminar nuclei of the thalamus. These are organized roughly by the area from which the projection arises. The thalamocortical network, which projects to the striatum, provides a wealth of inputs that represent a diverse array of signals related to representations of sensory inputs, motor programmes and internal states[2,5]. This dynamic set of inputs, which can change from moment to moment, therefore constitutes a generator of diversity. Moreover, the basal ganglia, and in particular the striatum, are capable of selection and differential amplification: in the short term through lateral inhibition and the membrane properties of the striatal projection neurons, which shift between different states

*Laboratory for Integrative Neuroscience, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, 5625 Fishers Lane, TS-13, Bethesda, Maryland 20892, USA. ‡Department of Psychology, Franz Hall, University of California, Los Angeles, California 10095-1563, USA.
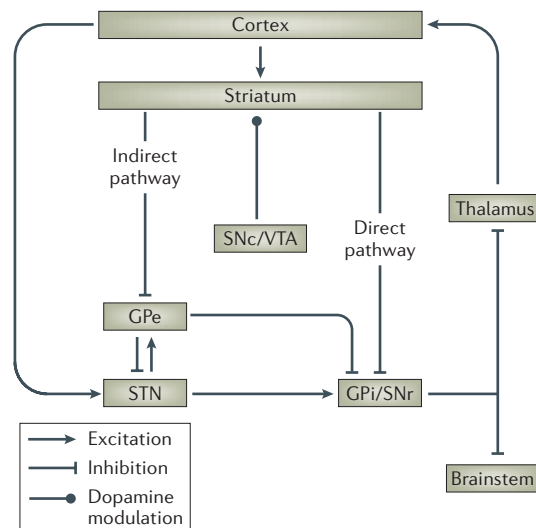Correspondence to B.J.K.
e-mail:
knowlton@psych.ucla.edu

Figure 1 | **A schematic of the main connections of the basal ganglia.** Simplified illustration of basal ganglia anatomy based on a primate brain. The direct and indirect pathways from the striatum have net effects of disinhibition and inhibition on the cortex, respectively. STN, subthalamic nucleus; GPe, external globus pallidus; GPi, internal globus pallidus; SNr, substantia nigra pars reticulta; SNc, substantia nigra pars compacta; VTA, ventral tegmental area.

of excitability; and in the long run by long-term synaptic plasticity, which can preserve or alter the process of behavioural selection[2,6].

Instrumental behaviour. Given their crucial place in the cerebrum, how do the basal ganglia function in generating purposive behaviour? Divac[7] and Konorski[8] were among the first to systematically examine the effects of cortical and basal ganglia lesions on the acquisition of instrumental behaviours. Whenever a particular outcome is contingent on a response, be it flexing a leg, traversing a maze or pressing a lever, the behaviour in question is instrumental. Instrumental behaviours differ from reflexes and fixed action patterns, which are not controlled by the contingency between behaviour and its consequences. Lesions of the sensorimotor cortex severely impaired skilled movements, and lesions of the premotor cortex impaired the chaining of action repertoires. By contrast, lesions of the basal ganglia (particularly the striatum) disrupted the very 'instrumentality' of actions — despite relatively intact fine movements, the animals that were tested could no longer perform or acquire actions in order to earn specific rewards or avoid aversive stimuli[8].

Although Konorski[8] presciently observed that striatal lesions produced variable results, he did not have at his disposal behavioural assays that would have allowed him to precisely analyse these effects. A major obstacle to understanding basal ganglia function is the conceptual confusion that characterized the field of instrumental learning for many decades, which in some ways persists

even today. Although instrumental behaviour appears to be primarily directed towards a goal, traditional theories, with a few notable exceptions[9,10], dismissed this obvious possibility. In the prime of behaviourism research, the study of learning was dominated by Hull and his followers, for whom instrumental learning is described in terms of stimulus–response (S–R) bonds strengthened by subsequent reinforcement[11,12]. S–R/reinforcement theory was based on the work of Thorndike, and aimed to eliminate 'unscientific' concepts such as intentionality, expectancy and internal representation[11,12]. The most fundamental assumption of this theory is that all behaviour is elicited by some antecedent stimuli from the external environment, and that the consequences of behaviour, by providing satisfaction or dissatisfaction to the organism, merely reinforces or weakens the S–R association. Deliberately dismissing the intentional account of goal-directed behaviour — that our behaviour can be controlled by action–outcome contingencies — the S–R/reinforcement theorist assigned no causal role to outcome expectancy. Although this position might be considered extreme today, its pervasive influence on neuroscience can hardly be exaggerated, and it remains powerful in many of the implicit assumptions made by researchers who interpret all neural activity solely as a function of antecedent stimuli presented before the motor response.

However, research over the past two decades has shown conclusively that animals can encode the causal relationship between their actions and outcomes, and control their actions according to their anticipation of, and desire for, the outcome[13,14]. Consequently, we are now aware of the paramount importance of two previously neglected variables — the remembered value of the expected outcome and the knowledge of the causal relationship between the action and the outcome. The realization that these variables can be manipulated by the experimenter has revolutionized the study of purposive behaviour.

As a result of this paradigm shift, there are now experimental assays that measure intentionality and goal-directedness. Two classes of assay have become common in the contemporary analysis of instrumental learning. In the first, the value of the outcome is increased (inflated) or decreased (devalued). Devaluation is far more common because it is easier to reduce the value of an outcome; for example, by giving the animal unlimited exposure to the food reinforcer before a brief probe test. If performance is sensitive to manipulations of outcome value (for example, if the rate of responding decreases after outcome devaluation), then the behaviour is controlled by the anticipation of the outcome. If performance is insensitive to these manipulations, then the behaviour is controlled by antecedent stimuli (it is habitual). Importantly, this test should occur in the absence of the outcome to probe the nature of memory for the association independently of new learning that can occur during the test.
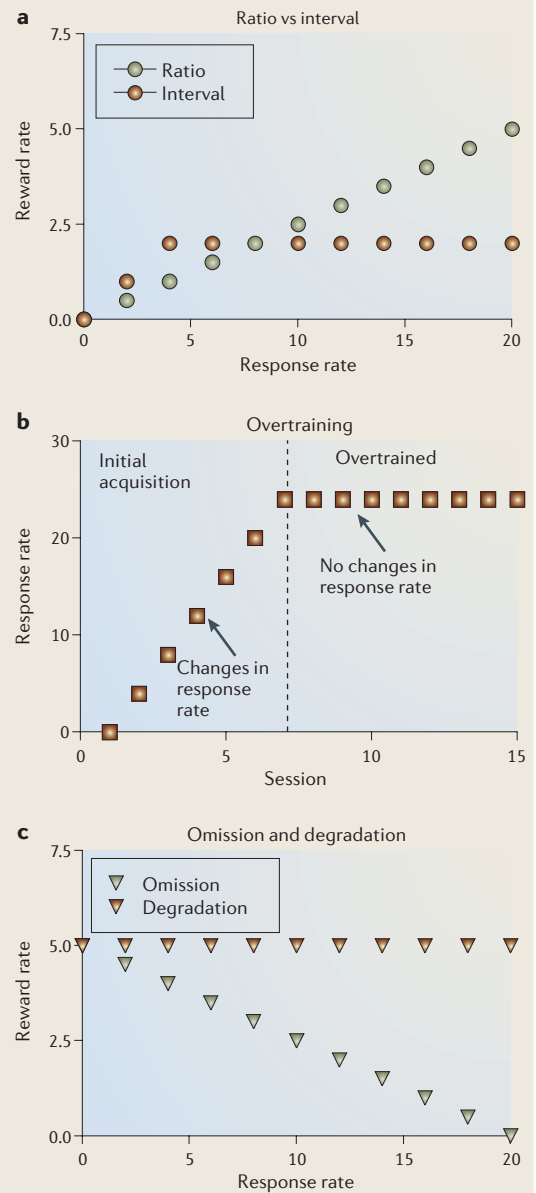
In the second class of assays, the action–outcome contingency (A–O; the degree to which the outcome depends on the action) is manipulated[14,15]. This is often

Box 1 | Conditions that lead to habit formation

In ratio schedules, a response results in a certain probability of reward; more responses yield more rewards. In interval schedules, a response is only rewarded after a certain time interval has elapsed. Under certain conditions (for example, when a single action–outcome (A–O) pairing is used), these schedules can generate behaviours that differ greatly in their sensitivity to manipulations of outcome value and instrumental contingency. For instance, training under an interval schedule results in behaviour that is less sensitive to the imposition of the omission contingency[17]. In short, whereas ratio schedules produce goal-directed actions controlled by the A–O contingency, interval schedules tend to generate stimulus–response (S–R) habits[25]. The most crucial difference between these schedules can be illustrated by plotting their feedback functions, with the rate of response on the x-axis and the rate of reward on the y-axis[106] (panel a). Whereas ratio schedules set up a strong correlation between response rates and reward rates, interval schedules do not[13].

Moreover, as Dickinson has observed, in both overtraining and interval schedules, the experienced instrumental contingency — the correlation between a change in response rate and a change in reward rate — is low[24,107]. In interval schedules, this experienced contingency is usually low. However, in ratio schedules the experienced contingency is high early in training, when response rates vary, resulting in varying local rates of reward; but with overtraining, the animals tend to respond at a consistently high rate, resulting in little change in the local rates of reward (panel b). Finally, this hypothesis also explains why, given two actions and two outcomes in training, behaviour was shown to be goal-directed even after extensive training with interval schedules[14], as this condition ensures that experienced contingency remains high (choosing one action would completely stop reward delivered by the other action).

The feedback function can also be used to illustrate common manipulations of the instrumental contingency. For example, omission is a complete reversal of the normal A–O contingency — that is, a response prevents the reinforcer, but no response results in reinforcer delivery (panel c). In degradation, the instrumental contingency is reduced by presenting non-contingent background reinforcers; for example, making the probability of the reinforcer the same regardless of response (panel c).

a | Ratio vs interval. Reward rate vs Response rate. Ratio, Interval.

b | Overtraining. Response rate vs Session. Initial acquisition, Overtrained. Changes in response rate. No changes in response rate.

c | Omission and degradation. Reward rate vs Response rate. Omission, Degradation.

done using contingency degradation, a procedure that introduces free rewards that are independent of any action. Instrumental contingency can be viewed as the probability of reward given a particular action relative to the probability of reward given no action. If these probabilities are the same, the contingency is said to be completely degraded. This would be the case, for example, if one is paid the same amount regardless of how much work is done; the question is to what extent work output would decrease as a result of the degraded contingency between work and pay. If degrading the contingency had no effect on work, it could be concluded that the behaviour was habitual and not goal-directed.

For any given behaviour to be established as a goal-directed action, it must pass both tests[16]. First, performance must be sensitive to revaluation of the outcome.

Second, performance must be sensitive to manipulation of the A–O contingency. Actions characterized by these criteria are not defined by specific motor programmes but by the goal state, such as a certain rate of reward; in maintaining this goal state the behaviour in question is modulated bidirectionally. Such bidirectional control can be demonstrated empirically by a complete reversal in instrumental contingency known as omission (BOX 1), in which an action that previously earned a reward is arranged to prevent it, and the animal can only earn rewards by refraining from performing the action[17,18]. Not surprisingly, omission is the most rapid method for reducing performance of goal-directed actions.

The analysis of the instrumental actions reviewed above has crucial implications for the study of habit formation, as behaviour not guided by outcome expectancy

and the instrumental contingency can be described as an S–R habit. This is a clear prediction from S–R/reinforcement theory, according to which the outcome is not part of the S–R association, but merely strengthens or weakens it. Indeed, under many conditions behaviours are not sensitive to changes in contingency and outcome devaluation[19–21]. The S–R/reinforcement theory of Thorndike and Hull has, therefore, stood the test of time when judged by its success at capturing the nature of habit learning.

As a result of extensive research, there is now a consensus that instrumental behaviours are controlled by two distinct systems — the A–O system and the S–R system — that are engaged under different conditions. In appetitive instrumental learning, the amount of training (in particular the number of rewarded responses) appears to be a crucial factor in determining the shift from A–O to S–R control over behaviour — that is, habit formation. Therefore, overtraining tends to promote habit formation[22]. The schedule of reinforcement used is also a key factor (BOX 1). Early studies using devaluation to examine the associative structure of instrumental conditioning failed to find any evidence that performance was controlled by goal expectancy, as devaluation had no effect on performance during the extinction test. The use of interval schedules in these studies was largely responsible for their failure to find evidence for A–O learning[19,23,24]. An explicit comparison of the schedules demonstrated that, even with the amount of reinforcement equated, interval schedules produce habitual responding whereas ratio schedules do not[25]. The difference in sensitivity to changes in outcome value must therefore be due to differences between interval and ratio schedules (BOX 1).

Habit learning in the dorsal striatum

Early efforts to understand basal ganglia functions were heavily influenced by S–R/reinforcement theory. According to the dominant view, the basal ganglia are the neural implementation of the law of effect, responsible for S–R learning reinforced by rewards (with the reinforcement signal possibly provided by dopamine) in a gradual process of habit formation[26–28]. Unsurprisingly, this view has initially found considerable empirical support[29,30].

Clear evidence comes from studies using the place/response learning task, first invented by Tolman and revived by Packard and McGaugh in a series of important experiments[31,32]. In this task, a rat is trained to retrieve food from one arm of a cross maze surrounded by various environmental cues (FIG. 2). After training, it is given probe tests in which the starting arm is placed at the opposite end of the maze. The use of the response strategy (same left turn) shows that the learning was inflexible and response-specific, but the use of the place strategy (right turn) shows that the animal was able to incorporate surrounding spatial cues in deciding which way to turn, selecting a response that was the opposite of what was initially learned.

After moderate training, most rats used the place strategy when tested, but after extensive training they switched to a response strategy. Moreover, with inactivation

of the dorsal striatum, the rats were more likely to use the place strategy despite extended training; however, inactivation of the hippocampus had the opposite effect — that is, the response strategy was used more frequently even early in training[32].

These results have two important implications. First, with overtraining, there is a shift in behavioural control from goal-directed actions to habits, and such a shift can be revealed by a behavioural assay. Second, the dorsal striatum and the hippocampus might, on the basis of this account, be viewed as competing learning systems. This view has been developed by Poldrack and Packard, who argued that direct or indirect neural connections between the hippocampus and dorsal striatum could mediate the competition between them[33].

Data from human studies suggest that there is a similar dissociation between declarative learning that is dependent on the medial temporal lobe (MTL) and non-declarative striatum-dependent learning. Unlike habits, declarative memories can be acquired rapidly, often after a single trial. These memories are explicit, in that participants are aware of the memories, and they are flexible, in that they can be applied to new situations. For example, declarative and habit learning were dissociated in a recent study using a concurrent discrimination task in which pairs of objects were presented[34]. The participant's task was to choose the rewarded item in each pair. Neurologically intact participants learned these discriminations quickly. Patients with severe amnesia following damage to the MTL were also able to learn these discriminations, but their performance improved much more slowly. Although the patients eventually learned the discriminations, they did not show explicit knowledge of these associations. They were unable to choose the rewarded items from the total array of stimuli. Their performance appeared to be habitual, with the presentation of the pair automatically eliciting the choice of the correct item. Indeed, the amnesic participants justified their choices by stating that some items "just seemed right", rather than relying on their declarative memory for previous trials.

Another task that has been used to assess habit learning in humans is the probabilistic classification task. In this task, a series of cues are each probabilistically associated with one of two outcomes, and the participant must guess which outcome is predicted on the basis of the cues that appear in each trial. Because the cues and outcomes are probabilistically associated, it is difficult to memorize their relationship explicitly. Amnesic patients are able to learn these associations normally, which is consistent with the idea that they are learned independently of MTL structures that support declarative memory. Furthermore, patients with Parkinson's disease, who exhibit abnormal striatal functioning due to loss of dopaminergic input, have been shown to be impaired in the implicit learning of these associations[35], although they managed to achieve normal levels of performance with further training. This suggests that other neural systems can support learning in this task. A recent study found that patients with mild Parkinson's disease were able to perform almost as well as neurologically

---

Extinction
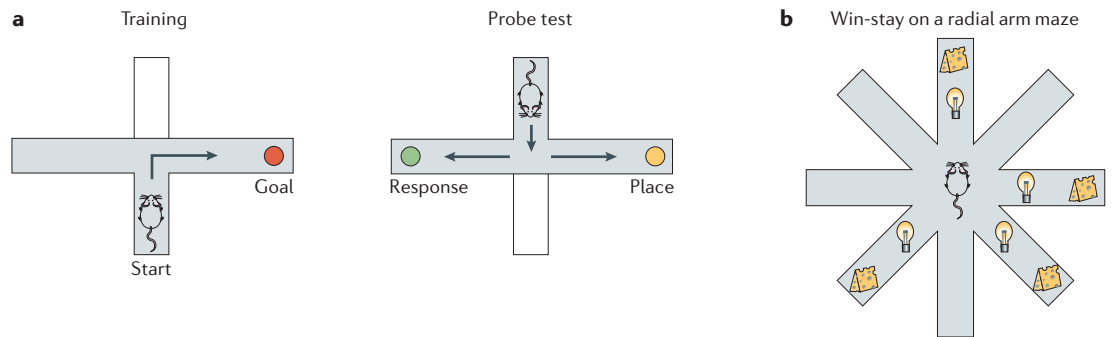Operationally, the withholding of reinforcement after previous reinforcement.

Figure 2 | **Simple maze tasks for measuring habits and actions. a** | In the place/response task, rats are trained to retrieve food from one arm of a T-maze or cross maze. The content of learning can be assessed by moving the starting arm to the other side of the maze on a probe test. The animal may enter the arm corresponding to the location of the reward during training (place strategy) or the arm corresponding to the turning response that was reinforced during training (response strategy). **b** | In the radial arm maze, animals can learn either a win-stay or a win-shift contingency. In win-stay, arms baited with food are signalled by a cue (such as a light at the entryway). Animals will gradually learn to respond to these cues by running down the arms and retrieving the food. Extensive win-stay training produces behaviour insensitive to devaluation[113], and requires the dorsolateral striatum[114]. By contrast, the win-shift task is similar to natural foraging in that animals need to efficiently traverse the region without revisiting areas before resources are replenished. They must learn the location of the arms that they have visited on each trial. Because arms are not re-baited on each trial, once the animal visits the arm and eats the food, it should remember not to return to that arm during the session. Win-shift performance is sensitive to devaluation[113], and is impaired by hippocampal lesions[114].

normal participants on the probabilistic classification task, but they showed a very different pattern of brain activation during performance as revealed by functional MRI. Whereas in control participants the striatal regions were activated during learning, patients with Parkinson's disease showed activation in the hippocampus and surrounding MTL cortical regions[36]. It appears that patients with Parkinson's disease achieved good performance by relying on declarative memory, whereas neurologically intact participants relied on non-declarative learning mechanisms. Many real-world tasks encountered by humans probably involve both habit and declarative learning; the system that contributes most to performance depends on the amount of training, the ease of memorizing associations and the relative integrity of the basal ganglia and MTL in the learner.

Functional heterogeneity in the dorsal striatum. Despite the evidence for basal ganglia involvement in habit learning, many findings cannot be explained by the idea that the dorsal striatum is the substrate of this type of learning. For example, studies recording from caudate cells in monkeys performing a saccade task have shown that the neural activity encoding the preferred direction of saccade could change according to whether that direction is rewarded, and this activity is rapidly modified as new contingencies are encountered[37–39]. Simultaneous recording from the prefrontal cortex (PFC) and caudate has shown that caudate activity rapidly adapts to the contingency before PFC activity does, and even before significant improvements in performance occur[40]. Such data suggest that certain learning mechanisms in the striatum do not have the characteristics of habit learning, that anticipation of future rewards has a crucial role in regulating striatal activity, and that changes in neural activity as a result of learning occur at a rate too rapid to

be explained by the slow and gradual changes posited by traditional S–R/reinforcement theory.

Because the dorsal striatum is a large and heterogeneous structure, similar to the cerebral cortex, the question naturally arises as to whether, like the cortex, it is also functionally specialized. The caudate in primates is part of the 'associative striatum', which receives inputs from association cortices. It corresponds to the dorsomedial striatum (DMS) in rodents, whereas the putamen is part of the sensorimotor striatum, corresponding to the dorsolateral striatum (DLS) in rodents[41] (FIG. 3a). Many investigators have created large lesions of the dorsal striatum in rodents, without regard for the medial/lateral distinction, but the damage appears to have been more prominent in the lateral region.

Indeed, the DLS differs from the DMS in connectivity, distribution of various receptors and mechanisms of synaptic plasticity[41–43] (BOX 2). Previous studies have also suggested a functional dissociation between the DLS and DMS[42,44]. For example, work by Devan and White showed that the DMS, like the dorsal hippocampus, is involved in flexible place learning, whereas the DLS subserves inflexible response learning[45,46]. In particular, they discovered that lesions of the DMS result in a preference for cue-based responding in the water-maze task. Taking into account these results and the different patterns of anatomical connectivity, these investigators proposed that the DMS belongs to the same functional system as the hippocampus.

In view of the distinction between actions and habits outlined above, these considerations raise the interesting possibility that the DLS is involved in S–R learning, whereas the DMS is involved in A–O learning. Yin *et al.* conducted a series of studies to test this hypothesis using assays (BOX 1) that could be
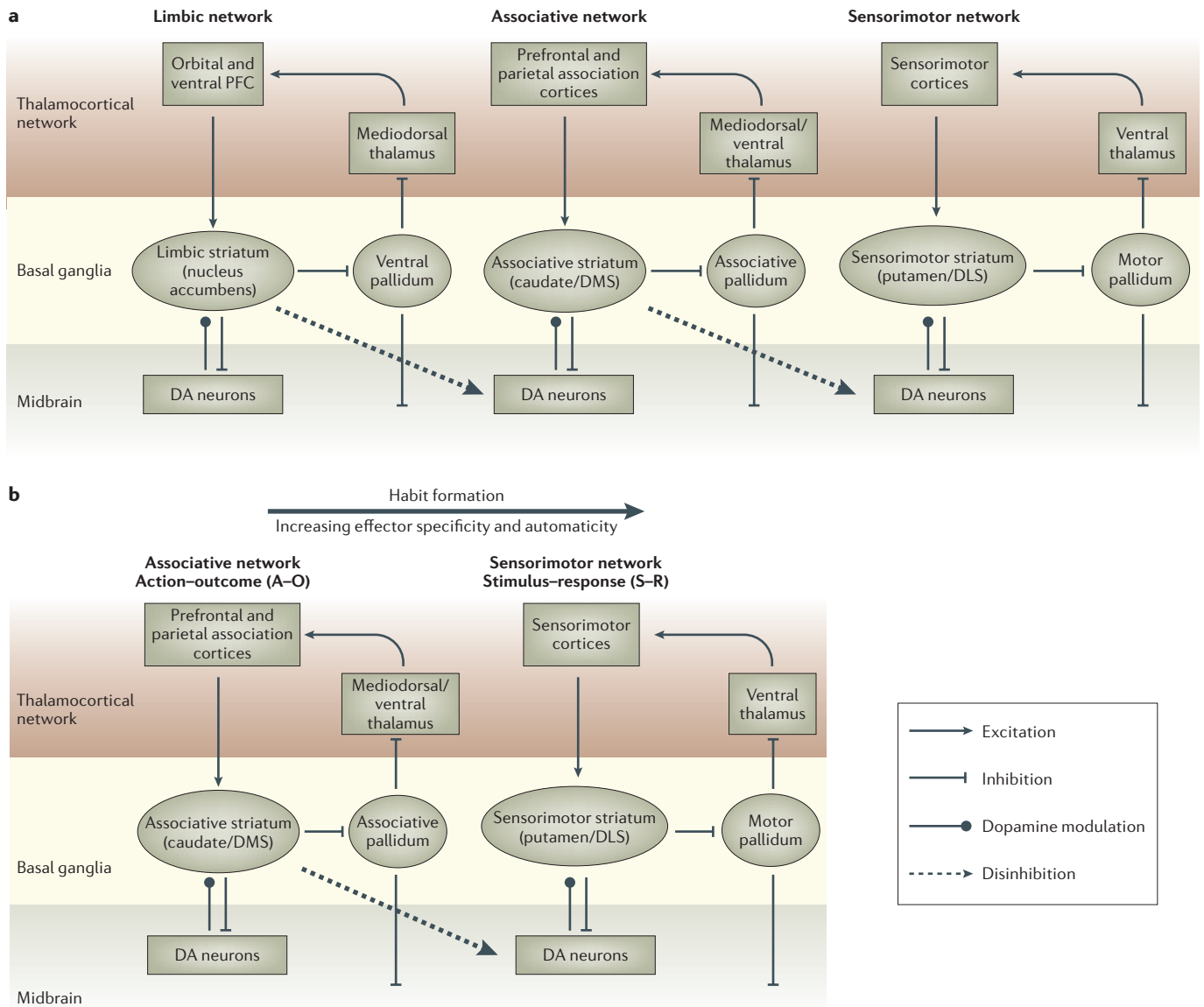
Figure 3 | **Cortico-basal ganglia networks as the fundamental motifs of cerebral organization. a** | Highly simplified schematic illustration of the three major networks: the limbic, associative and sensorimotor networks. **b** | Schematic illustration showing cortico-basal ganglia networks in relation to serial adaptation. A shift from the associative to the sensorimotor cortico-basal ganglia network is observed during habit formation. DA, dopamine; DLS, dorsolateral striatum; DMS, dorsomedial striatum; PFC, prefrontal cortex.

applied to instrumental learning paradigms[17,21,47–49]. Taking advantage of the established differences between ratio and interval feedback schedules, they first examined the effects of excitotoxic lesions to the DLS using variable interval schedules, which are known to generate habits — in this case, lever pressing that is insensitive to outcome devaluation. After training, the sucrose reward was devalued by inducing taste aversion until the animals stopped consuming it in their home cages. When these rats were tested later for extinction, lever pressing of controls was not reduced by devaluation. By contrast, although rats with DLS lesions could normally learn to press a lever for reward, they made fewer responses after devalu-

ation relative to the controls. It appears that because their habit system was disrupted by the lesion, the alternative A–O system assumed control over behaviour. However, a similar effect was not observed in rats with DMS lesions.

In another study, to assess the role of the DMS in A–O learning, Yin *et al.* used a training procedure with two actions and two outcomes under variable ratio schedules. This procedure generates goal-directed actions that are sensitive to outcome devaluation and contingency degradation[14]. The posterior DMS (pDMS) was shown to be a crucial substrate for the acquisition and expression of goal-directed actions. Both pre- and post-training lesions, as well as reversible inactivation of the pDMS

Box 2 | **Different rules of synaptic plasticity**

A basic assumption in contemporary neuroscience is that long-term synaptic plasticity, widely studied in the forms of long-term potentiation (LTP) and long-term depression (LTD), is a central physiological mechanism that underlies learning. The dissociation between the dorsolateral striatum (DLS) and dorsomedial striatum (DMS) at the level of behaviour is mirrored by distinct rules of synaptic plasticity in these regions. Although dopamine is crucial for all forms of striatal plasticity, the exact mechanisms show remarkable regional variation[43].

The DMS expresses LTP that depends on the activation of D1-like dopamine receptors and NMDA (*N*-methyl-D-aspartate) glutamate receptors[43,108]. The blockade of NMDA receptors in this region specifically prevents the learning of action–outcome contingency, suggesting a critical functional role for LTP in the DMS in such learning[47]. Additional evidence comes from a study using intracranial self-stimulation of the dopaminergic cells in rats to reinforce lever pressing[109]. The optimal parameters for self-stimulation were also found to induce cortico-striatal LTP *in vivo*, and the degree of potentiation in the cortico-striatal pathway in the DMS negatively correlated with the time taken to acquire lever pressing, which is a measure of initial action–outcome learning. This form of LTP requires the activation of D1 receptors, suggesting that it is the same form as is observed *in vitro*.

By contrast, dopamine-dependent striatal LTD is usually found in the DLS, and requires the activation of D2-like dopamine receptors, group I metabotropic glutamate receptors, and L-type calcium channels[110]. The resulting increase in intracellular calcium causes the postsynaptic synthesis and the release of endocannabinoids, which then act as a retrograde messenger on presynaptic cannabinoid CB1 receptors to decrease probability of glutamate release from the cortico-striatal terminals[111]. The role of this intriguing form of plasticity in DLS-dependent habit learning is not known. However, previous work has shown that local infusion of a D2 receptor agonist can improve acquisition on the win-stay task, which typically produces habitual responding that is insensitive to outcome devaluation[112,113].

As the above observations suggest, the same learning experience can result in different types of synaptic change in the DLS and DMS. These changes are regulated by distinct rules as a result of differential distribution of key receptors in these regions. Future studies will no doubt shed light on this striking correlation between mechanisms of striatal plasticity at the cellular and molecular level and the action/habit dissociation at the behavioural level.

abolished sensitivity to devaluation and degradation[48]. Moreover, local blockade of NMDA (*N*-methyl-D-aspartate) receptors, which are required for the induction of long-term potentiation (LTP) in this region, specifically prevented the encoding of new A–O contingencies without impairing performance[47]. Therefore, the pDMS appears to be a crucial neural substrate for the learning and expression of goal-directed actions. In its absence, the behaviour of the animal becomes habitual even under training conditions that result in goal-directed actions in control rats.

Moreover, it was shown that the pDMS is also involved in flexible choice behaviour in the place/response task on a cross maze[49]. After pre-training lesions were created, rats were trained extensively to retrieve food from the east arm of the maze, starting from the south arm, by turning right at the choice point (FIG. 2). Unlike control rats, most rats in the pDMS lesion group turned right on the probe tests, when they started from the north arm. This observation agrees with a growing body of recent data that show the role of the DMS in flexible choice behaviour[50,51]. Note that the key manipulation in the place/response task, namely the probe test with the opposite starting point, is similar to a reversal in the A–O contingency. Previously, a particular turn would lead to the arm with food, but with the 180° rotation of the starting point, the same turn would lead to the previously unrewarded arm. Again, the choice behaviour of rats with pDMS lesions is rendered inflexible and habitual.

Lever-pressing controlled by the instrumental contingency therefore shares common neural substrates with the use of the place strategy in the maze. Despite differences between the motor programmes of pressing a lever and of traversing a maze, the common neural

substrate in the pDMS suggests that this area is crucial for learning the A–O contingency, the feature shared by these tasks. On the cross maze, after a reversal in starting point, reaching the original goal requires a reintegration of the spatial features of the environment with the goal location. Whereas the hippocampus is necessary to ascertain the spatial location of the reward, the pDMS is involved in choosing the correct course of action that leads to this location.

One interpretation of these results is that the hippocampus does not compete with, or function independently of, the striatum, as has been previously claimed[29,33]. Rather, the hippocampus can act together with dorsomedial and ventral striatal regions to form a functional circuit. This hypothesis is supported by studies that examined activity in the DMS during spatial navigation on various mazes[52,53]. According to these studies, the DMS contains spatially selective neurons that fire when animals take a particular route to reach a goal; it also contains head-direction neurons with activity aligned with that of the place fields of hippocampal place cells. Therefore, information about the current position of the animal provided by hippocampal place cells can be used to signal where to go to reach a definite goal, and this information is probably conveyed to the DMS directly via the cortico-striatal projection from the hippocampal pyramidal neurons.

Further evidence for the role of the associative striatum in A–O learning has come from studies examining caudate (DMS homologue) activity in humans and other primates[54–56]. Tricomi *et al.*[57] found that caudate activity was modulated by the perceived contingency between action and outcome. Robust activation was found only when the participants thought that their action resulted in the gain or loss of money, whereas

time-locked anticipation of the outcome without the action contingency did not activate the caudate. These results also clearly implicate the associative striatum as a crucial component of the A–O system. Furthermore, Williams and Eskandar recorded neural activity from both the anterior caudate and the putamen (DLS homologue) in monkeys trained to move joysticks after presentations of discriminative stimuli[58]. These authors showed that caudate activity in response to outcome presentation is strongly correlated with the rate of learning (slope of the learning curve), whereas putamen activity is correlated with the learning curve itself. Although the authors interpreted such learning as S–R, in view of the framework above, the behaviour of the monkeys is probably controlled by the A–O contingency. The specific discriminative stimuli merely tells the animal which A–O contingency is in effect (that is, that a particular joystick movement will lead to reward), and the learning that occurs during the steepest portion of the learning curve corresponds to the initial acquisition of the A–O association. However, once this rapid learning has taken place, caudate activity quickly decreases, whereas putamen activity remains high and follows the learning curve closely until it asymptotes. This pattern of activity agrees with earlier theoretical claims about the relative rates of learning in the A–O and S–R systems[59]. Moreover, this study also found that, whereas stimulation of the putamen had no effect, stimulation of the caudate significantly enhanced the rate of learning without changing the asymptotic level of performance or hedonic preference, suggesting a causal role for this structure in instrumental learning.

A hierarchy of cortico-basal ganglia networks

We have suggested that associative structures — abstract descriptions of learning processes at the behavioural level — can be mapped onto discrete regions in the dorsal striatum. In particular, A–O learning can be mapped onto the DMS, whereas S–R learning can be mapped onto the DLS. How, then, are we to interpret such demonstrations of functional heterogeneity from studies that use the strategy of process dissociation? More importantly, what does it tell us about habit formation, whereby behavioural control is switched from one system to another?

Paradoxically, the chief implication of such functional heterogeneity is not that a more refined analysis of behaviour is accompanied by a more refined localization of function. If we compare the relevant data on the striatum with data from other brain regions that project to, or receive inputs from, the basal ganglia, a different picture emerges.

Considerable evidence shows that the PFC also has a crucial role in instrumental learning[60–65]. Studies by Balleine and colleagues have shown that rats with pre-training lesions to the medial PFC, especially the prelimbic region, which provides massive projections to the DMS, failed to show sensitivity to devaluation and degradation[60,61]. In addition, pre-training lesions of the mediodorsal nucleus of the thalamus, an eventual downstream target of outputs from the DMS as well as the major source of thalamic projections to the PFC, also

abolish sensitivity to devaluation and degradation[66]. To a certain extent, these observations resemble the effects of the pDMS lesions reviewed above[67].

Taking into account the above observations, we can no longer maintain that the dorsal striatum as a whole is a substrate for habit learning. Nor can we capture the distinction adequately with the traditional contrast between hippocampus-dependent learning and striatum-dependent learning. It should be noted that in this connection, selective pre-training lesions of the hippocampus do not consistently render behaviour habitual, as lesions of the pDMS do[68]. One possible role for the hippocampus, in view of this result and of the results from previous maze studies[31,46,69], is the integration of goal-directed actions that require some representation of spatial and/or temporal configurations. In any case, although the precise role of the hippocampus in A–O learning remains to be determined, the considerable functional heterogeneity in the dorsal striatum prompts a reconsideration of the currently accepted model of multiple memory systems in which the striatum as a whole serves a specific mnemonic function.

Alternatively, we propose that a cortico-basal ganglia network is a fundamental motif of cerebral organization, and is the fundamental unit of function at the level of behaviour (FIG. 3a). This claim is inspired by the traditional model of basal ganglia organization in terms of parallel and re-entrant loops[70], although we do not place special emphasis on either the thalamocortical target of basal ganglia outputs or the strictly parallel nature of the networks. Indeed, as discussed below, interaction between networks is vital to the transformation from actions to habits.

A cortico-basal ganglia network is a functional group comprising different cortical, striatal and pallidal components, in addition to the various cell groups (for example, dopaminergic) in the midbrain that constitute the brain's value system, as well as the associated diencephalic structures (for example, the thalamus and the subthalamic nucleus). The integration of various physiological processes in these components results in the output of the network — that is, behaviour. Although each of these components, by virtue of characteristic physiological properties, has unique 'computational' properties, at the behavioural level it is the integrated functioning of a distributed network comprising various components that is important. That is, when we probe behaviour with contemporary behavioural assays, we can map dissociable classes of behaviours onto dissociable cortico-basal ganglia networks. This point is worth emphasizing, as systems neuroscience is often dominated by attempts to localize psychological functions without regard for the actual functional circuitry of the brain. Not only do the psychological functions lack operational specificity, but the anatomical entities that are said to subserve such functions also lack the requisite circuitry. For instance, it is often asserted that the neocortex mediates a particular function, whereas the striatum subserves another[40]. By contrast, using operationally defined representational structures that can be dissociated behaviourally allows us to identify

the distributed networks that control distinct types of decision-making and learning. Although our proposal remains preliminary, and needs to be refined and corrected by future research, it should be clear by now that the traditional view of multiple memory systems (which divides the cerebrum into distinct functional systems corresponding to visually distinct anatomical entities such as the hippocampus, amygdala, striatum and neocortex) does not provide a fully satisfactory explanatory framework.

In the framework proposed here, the corticostriatal projections are loosely organized by cortical region so that the limbic cortex projects to the limbic striatum (mainly the nucleus accumbens), the association cortex projects to the dorsomedial, or associative, striatum, and the sensorimotor cortex projects to the dorsolateral or sensorimotor striatum[71] (FIG. 3a). The limbic network, which has a key role in appetitive Pavlovian learning, can exert tremendous influence on the associative and sensorimotor networks (discussed below).

In the associative network, the medial PFC (similar to the dorsolateral PFC in primates[72]) and the DMS (caudate) are involved in transient or 'working' memory. Lesions of either structure impair performance on spatial delayed response and delayed alternation tasks[7,73,74]. Like the PFC[62], caudate activity is also strongly modulated by anticipation of reward[75]. Thus, the associative network is capable of monitoring recent actions as well as anticipating their consequences. By contrast, the sensorimotor level comprises the sensorimotor cortices and their targets in the basal ganglia, beginning with the DLS. The outputs of this circuit eventually reach the motor cortices and brainstem motor networks. Unlike the associative striatum, neural activity in the sensorimotor striatum is not directly modulated by reward expectancy, but is more closely related to movements and to discriminative stimuli[76,77].

Habit formation and serial adaptation. Joel and Weiner proposed an important revision to the traditional scheme of parallel circuits[41,78,79]. Rather than closed loops with strict point-to-point topographical organization, they argued that interaction between different loops is made possible by interconnections between them. This claim is supported by recent anatomical work. In addition to the closed, strictly reciprocal projections, there are open striatonigral projections to a nigral area that, in turn, projects to a different striatal region[80]. These connections could allow the activity in one cortico-basal ganglia circuit to be propagated to the next circuit iteratively, suggesting a hierarchical organization in which a given cortico-basal ganglia circuit can be considered as a particular level in a functional hierarchy[81]. In addition, further interaction between circuits is possible at the level of the thalamo-cortico-thalamic connections[82].

We therefore propose that these overlapping cortico-basal ganglia networks form a labile hierarchy with three major levels, consisting of the limbic (stimulus–outcome, S–O), associative (A–O) and sensorimotor (S–R) networks (FIG. 3a). Here, we focus on the last two networks (FIG. 3b), which we locate in the two cortico-basal

ganglia circuits coursing through the dorsal striatum. These networks are characterized by strong re-entrant projections to the thalamocortical network, often precisely re-entering the cortical region from which the corticostriatal projections arise[83]. The associative network is crucial for the acquisition and performance of goal-directed actions, but in the course of habit formation this network appears to relinquish control over behaviour to the sensorimotor network, which is responsible for S–R habits. This relationship is most clearly revealed in two related sets of observations: one on differences in the extent of effector specificity, and the other on the switch, with extended practise, from one network to another in the control of behaviour.

Effector specificity refers to the extent to which the learning of a skill, as reflected in various performance measures, is limited to the effector (for example, a hand) with which it is originally trained. As shown by a study using monkeys, correct performance early in the learning of a behavioural sequence is not specific to the hand originally used to perform the sequence; with extensive practise, however, correct performance becomes specific to the hand used[84]. This task, not surprisingly, also requires the striatum, and learning of new and older sequences depends on different striatal regions.

The degree of effector specificity reflects the level of functional integration in the hierarchical organization of cortico-basal ganglia networks. The associative network achieves a higher level of functional integration, having at its disposal a wider range of motor programmes that can be selected to reach the goal. It is not effector-specific, possibly owing to the bilateral corticostriatal projections in this network. By contrast, the sensorimotor network is more effector-specific, possibly owing to its more lateralized corticostriatal projections[85]. With habit formation, therefore, the control of behaviour shifts from a higher level of functional integration to a lower one — more specifically, from the associative cortico-basal ganglia network to the sensorimotor cortico-basal ganglia network (FIG. 3b). However, extensive damage to either network results in the other network assuming control over instrumental behaviour[17,21,47–49,60].

Human imaging studies of habit learning have found that overtraining of a behaviour shifts the cortical substrate from ventral areas to more dorsal areas, and similar shifts have been observed in the striatum. Learning of new motor responses, for example, activated the caudate and the dorsolateral PFC, whereas with well-learned sequences the site of activation shifts to the putamen and motor cortices. When well-trained participants were asked to pay attention to their actions, the caudate and the more ventral PFC were again activated[86,87]. Such findings are not surprising in light of the hierarchical framework. Therefore, attention to action requires the associative network, but once a task is well learned only the sensorimotor network is needed for its performance.

In another study, Poldrack *et al.* examined the neural basis of automaticity, a concept from cognitive psychology operationally defined as resistance to interference from the performance of a secondary task[88]. After
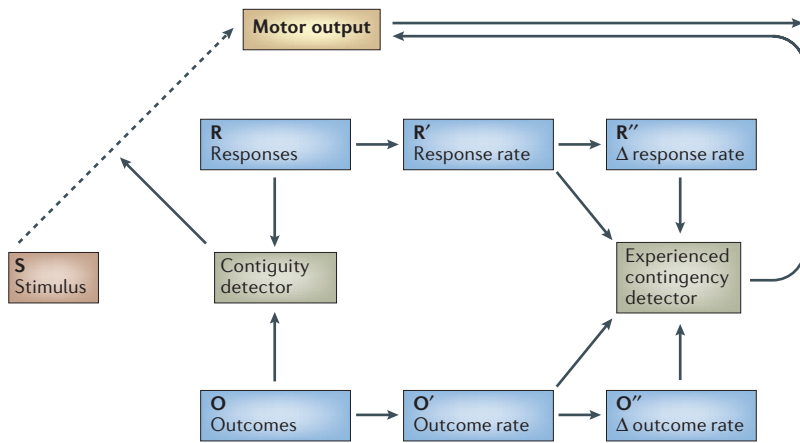
Figure 4 | **Schematic illustration of hypothetical mechanisms for the detection of instrumental contingency in appetitive instrumental learning.** The most straightforward mechanism for the detection of rates and changes in rates is the biological equivalent of differentiation. Anticipation is made possible by a higher-order derivative of the detected variable, just as velocity can increase more quickly than distance. Therefore, in the neural implementation of differentiation we already have a possible mechanism for prediction. The output of the 'experienced contingency detector' should have a crucial role in determining whether the action–outcome (A–O) system or the stimulus–response (S–R) system is controlling behaviour. In the absence of any activation of this detector, the S–R system, as described by traditional S–R/reinforcement theory, can assume control over behaviour. In this illustration we have also assumed that the contiguity between response and outcome reinforces the S–R association.

extensive training, the associative cortico-basal ganglia network, including the dorsolateral PFC and its corresponding striatal target in the caudate, decreased in activity. However, the supplementary motor area and the putamen/globus pallidus, parts of the sensorimotor cortico-basal ganglia network, did not show a similar decrease. As behaviour became more automatic with extensive practise, there was also a shift from the associative to the sensorimotor cortico-basal ganglia networks.

Potential mechanisms for serial adaptation. What are the mechanisms underlying the processes of serial adaptation described above? Unfortunately, there is little evidence available to answer this question. As mentioned above, the spiralling connections between the striatum and the midbrain discovered by Haber and colleagues could serve as a possible anatomical instantiation of links between networks, but numerous other possibilities exist[82]. Without indulging in speculative anatomy, we discuss the problem at a more abstract, computational level, which is open to different neural implementations.

As described in BOX 1, Dickinson first proposed that the experienced contingency between behaviour and reward is the key determinant of whether behaviour is goal-directed or habitual. Experienced contingency is defined as the correlation between changes in reward rates and changes in response rates. This account has implications for possible neural implementations. It suggests that there are neural detectors for rates of responses and rates of outcomes, and that outputs from these detectors must converge to yield some estimate of 'expe-

rienced contingency', which could determine whether the A–O system or the S–R system is engaged. To detect rates and changes in rates, a process akin to differentiation would be appropriate. For example, as illustrated by FIG. 4, activity in a particular unit could simply reflect the derivative (for example, rate) of activity upstream, and an iteration of this process could readily yield the second derivative (for example, a change in rate). Although our framework implicates the cortico-basal ganglia networks as the neural implementations of such computational processes, identifying the specific substrates requires extensive empirical work. This simple mechanism suggests that any reduction in experienced instrumental contingency, as encountered in contingency degradation and overtraining, could lead to reduced output of the contingency detector, and it is this output that would compete with the S–R/reinforcement system for the control of behaviour.

A different and more formal model, which accounts for much of the data on the various conditions leading to habit formation, was provided by a recent theoretical paper[89]. Using a set of computational methods known as reinforcement learning, Daw *et al.* modelled the process of habit formation by combining two independent controllers with distinct mechanisms for estimating value functions (the 'yield' of behaviour in a given state). The 'model-based' controller was used to simulate the A–O system, whereas the 'model-free' controller was used to simulate the S–R habit system. The key proposal was that arbitration is based on the uncertainty (posterior variances of estimated values or expected inaccuracy) in estimating the value function; the value determining actual choice behaviour is taken from the controller with the least uncertainty. According to Daw *et al.,* the model-free (habit) controller, using the temporal-difference algorithm, estimates value functions by caching — that is, storing a long-run value for future use — and choice behaviour is determined by the stored value. Because such estimates are divorced from the outcome (much like the S–R reinforcement theory), this method is computationally tractable but inflexible, yielding behaviour that is insensitive to outcome devaluation, whereas exactly the opposite is true of the model-based controller (A–O system). Further work is needed to extend the uncertainty-based model beyond discrete Markov decision processes to truly free operant conditions, and to incorporate instrumental contingency into this model.

Habits in relation to addiction

Addiction has often been viewed simply as a maladaptive type of habit learning[90]. Although this view is supported by the insensitivity of drug-seeking behaviour to harmful consequences, the motivational compulsion seen in addiction can hardly be explained by S–R/reinforcement theory alone. Although our suggestion that habit formation involves the serial adaptation of distinct cortico-basal ganglia networks is also supported by the literature on addiction[91,92], in the case of addiction considerations must be given to additional processes, especially appetitive Pavlovian conditioning, as incidental pairing between situational cues and drugs allow such learning

Temporal-difference algorithm
A reinforcement learning method that is driven by the difference between temporally successive predictions, rather than by the difference between predicted and actual outcomes.

Markov decision processes
A stochastic control process with the Markov property: future states are conditionally independent of past states and depend only on the current state.

to take place. In most situations, of course, Pavlovian conditioning and instrumental learning can occur simultaneously, and interact in controlling behaviour. In our view, to understand addiction it is necessary to consider these interactions.

In Pavlovian conditioning, the contingent pairing of a conditional stimulus (CS) and an outcome results in the acquisition of conditional responses (CRs) to the previously neutral stimulus. The CR is not controlled by the response–outcome contingency: even if the response prevents the outcome, as when an omission contingency (BOX 1) is imposed, the CR is still elicited by the CS[93].

As Berridge and Robinson have argued, situational cues in addiction can acquire motivational properties, which they call 'incentive salience'[94]. Incentive salience is a measure of how much the reward is 'wanted' rather than 'liked', and it is this property that is argued to be greatly enhanced in addiction. Being a description of appetitive preparatory CRs, it can be dissociated from consummatory CRs such as taste reactivity[8,95]. Preparatory CRs are usually less specific than consummatory CRs (for example, salivation); although measurable peripherally, they also correspond to central motivational states such as craving or wanting in appetitive learning, or fear in aversive learning[8]. Such states induced by predictors of reward can directly potentiate instrumental responding[8,96].

It has long been claimed that discriminative stimuli preceding instrumental actions and reafferent stimuli generated by actions can form associations with the outcome and further motivate instrumental behaviour[97]. Although such explanations fail to account for much of the contemporary data, they remain valuable for their emphasis on Pavlovian–instrumental interactions, which have been amply documented[24]. Pavlovian–instrumental transfer (PIT), a rigorous experimental method used to study such interactions, assesses the extent to which Pavlovian CSs that predict outcomes can potentiate instrumental performance yielding the same outcomes[98]. As PIT is normally produced by long, tonic CSs, which can also elicit preparatory CRs in appetitive conditioning, one potentially important mechanism underlying addiction at the level of neural systems is the heightened transfer from the Pavlovian incentive system to the systems that govern instrumental behaviour[8]. This mechanism is in accord with the important role of environmental cues in triggering compulsive drug seeking[91].

In view of the serial adaptation hypothesis described above, PIT can also be viewed in terms of interactions between cortico-basal ganglia networks (FIG. 3a). In this connection, an intriguing recent finding is that as behaviour becomes habitual it also becomes more susceptible to transfer of control — that is, a Pavlovian CS can potentiate habitual responding more than it can potentiate goal-directed actions[99]. As the nucleus accumbens, which belongs to the limbic cortico-basal ganglia network, is critical for PIT[100], it could also exert control over the sensorimotor network (FIG. 3a) via the spiralling connections with dopaminergic neurons[80].

Similar ideas have been advanced recently by Canales, whose argument is based on experiments that measure activity in different chemical compartments in the striatum[5]. Work by Canales and Graybiel has shown that exposure to addictive drugs leads to relatively higher activation of striosomal neurons than of matrix neurons, and that this pattern of activation is correlated with a measure of motor stereotypy[101]. These two compartments generally delineate two sources of cortical inputs to the striatum, and so Canales argues that the dominance of the striosomal activation reflects heightened control of the basal ganglia circuitry by inputs from limbic cortical areas. This hypothesis is supported by the finding that lesioning or inactivating the infralimbic cortex, which is involved in the inhibitory control of Pavlovian CRs[102] and a source of inputs to the striosome compartment, resulted in sensitivity to devaluation even in overtrained rats whose performance is normally habitually controlled[103,104]. Although the role of the infralimbic–striosome system in habit formation is not clear, it may in fact be engaged in Pavlovian control of instrumental systems. An obvious prediction here is that lesions of this system would disrupt PIT.

What is clear from the above discussion is that the motivational compulsion seen in addiction could be modelled by PIT, and implemented by links between the limbic and the sensorimotor cortico-basal ganglia networks (FIG. 3a). Accordingly, different stages of addiction are expected to be characterized by distinct behavioural characteristics as a result of the underlying serial adaptation from network to network. In support of such claims, a recent study of the effect of cocaine self-administration on striatal activity in monkeys found a gradual spread and intensification of the effects of the drug from the ventral striatum to the dorsal striatum[92]. Everitt and Robbins have also shown that reafferent stimuli that predict reward can initially potentiate dopamine release in the accumbens, and eventually in the dorsal striatum, which suggests that these Pavlovian motivators can affect cortico-basal ganglia networks that mediate instrumental behaviour[105]. Pavlovian learning, therefore, possibly precedes instrumental learning, with serial adaptation initiated in the limbic network and eventually spreading to the sensorimotor network. As a result, our general framework can readily incorporate various accounts of addiction, and establish a relationship between habitual responding and motivational compulsion.

Conclusions

Given the enormous structural complexity of the basal ganglia, a strictly bottom-up approach in elucidating their functions might not be fruitful. Instead, research can be guided by a top-down analysis based on the understanding of behaviour. The goal of this review, above all, is to clear up conceptual confusions and stimulate research by outlining a coherent framework based on known anatomy and physiology as well as our current understanding of instrumental behaviours.

Central to this framework is the distinction between goal-directed actions and stimulus-driven habits, the two main categories of instrumental behaviour. They can be dissociated at the behavioural level using assays that manipulate the value of the outcome and the contingency between action and outcome. Using these assays,

---

Stereotypy
Repetitive patterns of behaviour that are characterized by the lack of variation; often observed in various psychiatric disorders and after psychomotor stimulant administration.

Striosome
A patch-like compartment in the striatum that is characterized by low acetylcholinesterase staining and other chemical markers.

they can also be dissociated in terms of their underlying neural substrates, in the form of distinct cortico-basal ganglia networks.

Clearly, an understanding of network interactions that result in a switch in behavioural control from actions to habits has important implications for the study of skill learning, addiction and various clinical disorders resulting from basal ganglia abnormalities. At present, however, we remain ignorant of the detailed mechanisms that underlie habit formation at all levels of analysis. At the behavioural level, all the conditions that promote habit formation have yet to be characterized precisely. Although several behavioural characteristics of habits can be specified (for example, insensitivity to outcome devaluation and contingency degradation, lack of behavioural flexibility and lack of awareness in humans), other characteristics are less clear (for example, the degree of

effector specificity and the need for attention during learning). At the neural systems level, we do not yet understand the properties of the cortico-basal ganglia networks responsible for differences in behavioural flexibility, or in sensitivity to instrumental contingency manipulations. At the cellular level, in addition to our ignorance of the detailed molecular mechanisms underlying synaptic transmission and plasticity in the basal ganglia, we do not yet understand how synaptic plasticity in the basal ganglia alters the outputs of the networks, and we do not have direct evidence linking such plasticity to well-defined learning. Nevertheless, we hope that the framework proposed here will stimulate future research, by directing attention to those variables that are crucial in the analysis of purposive behaviour, and by underscoring the importance of precise behavioural analysis in elucidating the functions of neural systems.

1. Swanson, L. W. Cerebral hemisphere regulation of motivated behavior. *Brain Res.* **886**, 113–164 (2000).
A learned and provocative review of cerebral anatomy focusing on basal ganglia organization.
2. Wilson, C. J. in *The Synaptic Organization of the Brain* (ed. Shepherd, G. M.) 329–375 (Oxford Univ. Press, New York, 2004).
3. Deniau, J. M. & Chevalier, G. Disinhibition as a basic process in the expression of striatal functions. II. The striato-nigral influence on thalamocortical cells of the ventromedial thalamic nucleus. *Brain Res.* **334**, 227–233 (1985).
4. Albin, R. L., Young, A. B. & Penney, J. B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12**, 366–375 (1989).
5. Canales, J. J. Stimulant-induced adaptations in neostriatal matrix and striosome systems: transiting from instrumental responding to habitual behavior in drug addiction. *Neurobiol. Learn. Mem.* **83**, 93–103 (2005).
6. Wickens, J. R. & Koetter, R. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.)187–214 (MIT Press, Cambridge, Massachusetts, 1995).
7. Divac, I., Rosvold, H. E. & Szwarbart, M. K. Behavioral effects of selective ablation of the caudate nucleus. *J. Comp. Physiol. Psychol.* **63**, 184–190 (1967).
8. Konorski, J. *Integrative Activity of the Brain* (University of Chicago Press, Chicago, 1967).
9. Skinner, B. *The Behavior of Organisms* (Appleton-Century-Crofts, New York, 1938).
10. Tolman, E. C. *Purposive Behavior in Animals and Man* (Macmillan, New York, 1932).
11. Thorndike, E. L. *Animal Intelligence: Experimental Studies* (Macmillan, New York, 1911).
12. Hull, C. *Principles of Behavior* (Appleton-Century-Crofts, New York, 1943).
13. Dickinson, A. in *Animal Learning and Cognition* (ed. Mackintosh, N. J.) 45–79 (Academic, Orlando, 1994).
14. Colwill, R. M. & Rescorla, R. A. in *The Psychology of Learning and Motivation* (ed. Bower, G.) 55–104 (Academic, New York, 1986).
References 13 and 14 are excellent introductions to the modern study of instrumental learning.
15. Hammond, L. J. The effect of contingency upon the appetitive conditioning of free-operant behavior. *J. Exp. Anal. Behav.* **34**, 297–304 (1980).
16. Dickinson, A. & Balleine, B. in *Spatial Representation: Problems in Philosophy and Psychology* (eds Eilan, N. et al.) 277–293 (Blackwell, Malden, Massachusetts, 1993).
17. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behav. Brain Res.* **166**, 189–196 (2006).
18. Davis, J. & Bitterman, M. E. Differential reinforcement of other behavior (DRO): a yoked-control comparison. *J. Exp. Anal. Behav.* **15**, 237–241 (1971).
19. Holman, E. W. Some conditions for the dissociation of consummatory and instrumental behavior in rats. *Learn. Motiv.* **6**, 358–366 (1975).

20. Adams, C. D. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol.* **33B**, 109–122 (1982).
21. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* **19**, 181–189 (2004).
22. Colwill, R., Rescorla, R. A. The role of response–reinforcer associations increases throughout extended instrumental training. *Anim. Learn. Behav.* **16**, 105–111 (1988).
23. Dickinson, A. in *Learning, Motivation, and Cognition* (eds Bouton, M. E. & Fanselow, M. S.) 345–367 (American Psychological Association, Washington DC, 1997).
24. Dickinson, A. in *Contemporary Learning Theories* (eds Klein, S. B. & Mowrer, R. R.) 279–308 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1989).
25. Dickinson, A., Nicholas, D. J. & Adams, C. D. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q. J. Exp. Psychol. B* **35**, 35–51 (1983).
26. Miller, R. *Meaning and Purpose in the Intact Brain* (Oxford Univ. Press, New York, 1981).
27. Mishkin, M., Malamut, B. & Bachevalier, J. in *Neurobiology of Learning and Memory* (eds Lynch, G. et al.) 65–77 (Guilford, New York, 1984).
28. Robbins, T. W., Giardini, V., Jones, G. H., Reading, P. & Sahakian, B. J. Effects of dopamine depletion from the caudate-putamen and nucleus accumbens septi on the acquisition and performance of a conditional discrimination task. *Behav. Brain Res.* **38**, 243–261 (1990).
29. Packard, M. G. & Knowlton, B. J. Learning and memory functions of the basal ganglia. *Annu. Rev. Neurosci.* **25**, 563–593 (2002).
30. White, N. M. A functional hypothesis concerning the striatal matrix and patches: mediation of S–R memory and reward. *Life Sci.* **45**, 1943–1957 (1989).
31. Packard, M. G. Glutamate infused posttraining into the hippocampus or caudate-putamen differentially strengthens place and response learning. *Proc. Natl Acad. Sci. USA* **96**, 12881–12886 (1999).
32. Packard, M. G. & McGaugh, J. L. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol. Learn. Mem.* **65**, 65–72 (1996).
33. Poldrack, R. A. & Packard, M. G. Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* **41**, 245–251 (2003).
34. Bayley, P. J., Frascino, J. C. & Squire, L. R. Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature* **436**, 550–553 (2005).
35. Knowlton, B. J., Mangels, J. A. & Squire, L. R. A neostriatal habit learning system in humans. *Science* **273**, 1399–1402 (1996).
36. Moody, T. D., Bookheimer, S. Y., Vanek, Z. & Knowlton, B. J. An implicit learning task activates medial temporal lobe in patients with Parkinson's disease. *Behav. Neurosci.* **118**, 438–442 (2004).

37. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neurosci.* **1**, 411–416 (1998).
38. Lauwereyns, J. et al. Feature-based anticipation of cues that predict reward in monkey caudate nucleus. *Neuron* **33**, 463–473 (2002).
39. Lauwereyns, J., Watanabe, K., Coe, B. & Hikosaka, O. A neural correlate of response bias in monkey caudate nucleus. *Nature* **418**, 413–417 (2002).
40. Pasupathy, A. & Miller, E. K. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**, 873–876 (2005).
41. Joel, D. & Weiner, I. The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* **96**, 451–474 (2000).
42. West, M. O. et al. A region in the dorsolateral striatum of the rat exhibiting single-unit correlations with specific locomotor limb movements. *J. Neurophysiol.* **64**, 1233–1246 (1990).
43. Partridge, J. G., Tang, K. C. & Lovinger, D. M. Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J. Neurophysiol.* **84**, 1422–1429 (2000).
The first study to demonstrate regional variations in the types and mechanisms of striatal synaptic plasticity.
44. Whishaw, I. Q., Mittleman, G., Bunch, S. T. & Dunnett, S. B. Impairments in the acquisition, retention and selection of spatial navigation strategies after medial caudate-putamen lesions in rats. *Behav. Brain Res.* **24**, 125–138 (1987).
45. Devan, B. D., McDonald, R. J. & White, N. M. Effects of medial and lateral caudate-putamen lesions on place- and cue-guided behaviors in the water maze: relation to thigmotaxis. *Behav. Brain Res.* **100**, 5–14 (1999).
46. Devan, B. D. & White, N. M. Parallel information processing in the dorsal striatum: relation to hippocampal function. *J. Neurosci.* **19**, 2789–2798 (1999).
47. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *Eur. J. Neurosci.* **22**, 505–512 (2005).
48. Yin, H. H., Ostlund, S. B., Knowlton, B. J. & Balleine, B. W. The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* **22**, 513–523 (2005).
49. Yin, H. H. & Knowlton, B. J. Contributions of striatal subregions to place and response learning. *Learn. Mem.* **11**, 459–463 (2004).
References 47–49 present a series of studies that established for the first time a dissociation between S–R learning in the DLS and A–O learning in the pDMS.
50. Ragozzino, M. E. Acetylcholine actions in the dorsomedial striatum support the flexible shifting of response patterns. *Neurobiol. Learn. Mem.* **80**, 257–267 (2003).

51. Ragozzino, M. E., Jih, J. & Tzavos, A. Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res.* **953**, 205–214 (2002).

52. Ragozzino, K. E., Leutgeb, S. & Mizumori, S. J. Dorsal striatal head direction and hippocampal place representations during spatial navigation. *Exp. Brain Res.* **139**, 372–376 (2001).

53. Mulder, A. B., Tabuchi, E. & Wiener, S. I. Neurons in hippocampal afferent zones of rat striatum parse routes into multi-pace segments during maze navigation. *Eur. J. Neurosci.* **19**, 1923–1932 (2004).

54. Delgado, M. R., Locke, H. M., Stenger, V. A. & Fiez, J. A. Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn. Affect. Behav. Neurosci.* **3**, 27–38 (2003).

55. Delgado, M. R., Stenger, V. A. & Fiez, J. A. Motivation-dependent responses in the human caudate nucleus. *Cereb. Cortex* **14**, 1022–1030 (2004).

56. Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C. & Berns, G. S. Human striatal responses to monetary reward depend on saliency. *Neuron* **42**, 509–517 (2004).

57. Tricomi, E. M., Delgado, M. R. & Fiez, J. A. Modulation of caudate activity by action contingency. *Neuron* **41**, 281–292 (2004).
An interesting human imaging study that provided strong evidence for the role of the caudate in encoding A–O contingencies.

58. Williams, Z. M. & Eskandar, E. N. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nature Neurosci.* **9**, 562–568 (2006).

59. Dickinson, A., Balleine, B., Watt, A. & Gonzalez, F. Motivational control after extended instrumental training. *Anim. Learn. Behav.* **23**, 197–206 (1995).

60. Balleine, B. W. & Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407–419 (1998).

61. Corbit, L. H. & Balleine, B. W. The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res.* **146**, 145–157 (2003).

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

63. Tsujimoto, S. & Sawaguchi, T. Properties of delay-period neuronal activity in the primate prefrontal cortex during memory- and sensory-guided saccade tasks. *Eur. J. Neurosci.* **19**, 447–457 (2004).

64. Tsujimoto, S. & Sawaguchi, T. Neuronal representation of response–outcome in the primate prefrontal cortex. *Cereb. Cortex* **14**, 47–55 (2004).

65. Tsujimoto, S. & Sawaguchi, T. Working memory of action: a comparative study of ability to selecting response based on previous action in New World monkeys (*Saimiri sciureus* and *Callithrix jacchus*). *Behav. Processes* **58**, 149–155 (2002).

66. Corbit, L. H., Muir, J. L. & Balleine, B. W. Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur. J. Neurosci.* **18**, 1286–1294 (2003).

67. Ostlund, S. B. & Balleine, B. W. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.* **25**, 7763–7770 (2005).

68. Corbit, L. H., Ostlund, S. B. & Balleine, B. W. Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J. Neurosci.* **22**, 10976–10984 (2002).

69. Packard, M. G. & McGaugh, J. L. Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: further evidence for multiple memory systems. *Behav. Neurosci.* **106**, 439–446 (1992).

70. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).

71. Reep, R. L., Cheatwood, J. L. & Corwin, J. V. The associative striatum: organization of cortical projections to the dorsocentral striatum in rats. *J. Comp. Neurol.* **467**, 271–292 (2003).

72. Dalley, J. W., Cardinal, R. N. & Robbins, T. W. Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neurosci. Biobehav. Rev.* **28**, 771–784 (2004).

73. Divac, I., Markowitsch, H. J. & Pritzel, M. Behavioral and anatomical consequences of small intrastriatal injections of kainic acid in the rat. *Brain Res.* **151**, 523–532 (1978).

74. Levy, R., Friedman, H. R., Davachi, L. & Goldman-Rakic, P. S. Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *J. Neurosci.* **17**, 3870–3882 (1997).

75. Hassani, O. K., Cromwell, H. C. & Schultz, W. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* **85**, 2477–2489 (2001).

76. Kimura, M., Aosaki, T. & Ishida, A. Neurophysiological aspects of the differential roles of the putamen and caudate nucleus in voluntary movement. *Adv. Neurol.* **60**, 62–70 (1993).

77. Kanazawa, I., Murata, M. & Kimura, M. Roles of dopamine and its receptors in generation of choreic movements. *Adv. Neurol.* **60**, 107–112 (1993).

78. Joel, D. & Weiner, I. The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience* **63**, 363–379 (1994).
An important review in a series by the same authors arguing for interactions between cortico-basal ganglia networks.

79. Joel, D. & Weiner, I. The connections of the primate subthalamic nucleus: indirect pathways and the open-interconnected scheme of basal ganglia-thalamocortical circuitry. *Brain Res. Brain Res. Rev.* **23**, 62–78 (1997).

80. Haber, S. N., Fudge, J. L. & McFarland, N. R. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J. Neurosci.* **20**, 2369–2382 (2000).

81. Redgrave, P., Prescott, T. J. & Gurney, K. The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* **89**, 1009–1023 (1999).

82. Haber, S. N. The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.* **26**, 317–330 (2003).

83. Middleton, F. A. & Strick, P. L. Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Res. Brain Res. Rev.* **31**, 236–250 (2000).

84. Rand, M. K. *et al.* Characteristics of sequential movements during early learning period in monkeys. *Exp. Brain Res.* **131**, 293–304 (2000).

85. McGeorge, A. J. & Faull, R. L. The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* **29**, 503–537 (1989).

86. Jueptner, M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. & Passingham, R. E. Anatomy of motor learning. II. Subcortical structures and learning by trial and error. *J. Neurophysiol.* **77**, 1325–1337 (1997).

87. Jueptner, M. *et al.* Anatomy of motor learning. I. Frontal cortex and attention to action. *J. Neurophysiol.* **77**, 1313–1324 (1997).

88. Poldrack, R. A. *et al.* The neural correlates of motor skill automaticity. *J. Neurosci.* **25**, 5356–5364 (2005).
References 85–88 show shifts in activation patterns of cortico-basal ganglia networks in the course of skill learning.

89. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neurosci.* **8**, 1704–1711 (2005).

90. Everitt, B. J. & Wolf, M. E. Psychomotor stimulant addiction: a neural systems perspective. *J. Neurosci.* **22**, 3312–3320 (2002).

91. Altman, J. *et al.* The biological, social and clinical bases of drug addiction: commentary and debate. *Psychopharmacology (Berl.)* **125**, 285–345 (1996).

92. Porrino, L. J., Lyons, D., Smith, H. R., Daunais, J. B. & Nader, M. A. Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. *J. Neurosci.* **24**, 3554–3562 (2004).

93. Williams, D. R. & Williams, H. Automaintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J. Exp. Anal. Behav.* **12**, 511–520 (1969).

94. Robinson, T. E. & Berridge, K. C. Addiction. *Annu. Rev. Psychol.* **54**, 25–53 (2003).

95. Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.* **28**, 309–369 (1998).

96. Tiffany, S. T. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychol. Rev.* **97**, 147–168 (1990).

97. Rescorla, R. A. & Solomon, R. L. Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* **74**, 151–182 (1967).

98. Corbit, L. H. & Balleine, B. W. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian–instrumental transfer. *J. Neurosci.* **25**, 962–970 (2005).

99. Holland, P. C. Relations between Pavlovian–instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process.* **30**, 104–117 (2004).

100. Corbit, L. H., Muir, J. L. & Balleine, B. W. The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* **21**, 3251–3260 (2001).

101. Canales, J. J. & Graybiel, A. M. A measure of striatal function predicts motor stereotypy. *Nature Neurosci.* **3**, 377–383 (2000).

102. Rhodes, S. E. & Killcross, S. Lesions of rat infralimbic cortex enhance recovery and reinstatement of an appetitive Pavlovian response. *Learn. Mem.* **11**, 611–616 (2004).

103. Coutureau, E. & Killcross, S. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* **146**, 167–174 (2003).

104. Killcross, S. & Coutureau, E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* **13**, 400–408 (2003).

105. Everitt, B. J. & Robbins, T. W. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neurosci.* **8**, 1481–1489 (2005).

106. Baum, W. M. The correlation-based law of effect. *J. Exp. Anal. Behav.* **20**, 137–153 (1973).

107. Dickinson, A. Actions and habits: the development of behavioural autonomy. *Phil. Trans. R. Soc. Lond. B* **308**, 67–78 (1985).

108. Kerr, J. N. & Wickens, J. R. Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum *in vitro*. *J. Neurophysiol.* **85**, 117–124 (2001).

109. Reynolds, J. N., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. *Nature* **413**, 67–70 (2001).

110. Gerdeman, G. L., Partridge, J. G., Lupica, C. R. & Lovinger, D. M. It could be habit forming: drugs of abuse and striatal synaptic plasticity. *Trends Neurosci.* **26**, 184–192 (2003).

111. Gerdeman, G. L., Ronesi, J. & Lovinger, D. M. Postsynaptic endocannabinoid release is critical to long-term depression in the striatum. *Nature Neurosci.* **5**, 446–451 (2002).

112. Packard, M. G. & White, N. M. Dissociation of hippocampus and caudate nucleus memory systems by posttraining intracerebral injection of dopamine agonists. *Behav. Neurosci.* **105**, 295–306 (1991).

113. Sage, J. R. & Knowlton, B. J. Effects of US devaluation on win-stay and win-shift radial maze performance in rats. *Behav. Neurosci.* **114**, 295–306 (2000).

114. Packard, M. G., Hirsh, R. & White, N. M. Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J. Neurosci.* **9**, 1465–1472 (1989).

DATABASES
The following terms in this article are linked online to:
OMIM: http://www.ncbi.nlm.nih.gov/Omim
Parkinson's disease

FURTHER INFORMATION
Knowlton's homepage: http://www.psych.ucla.edu/Faculty/Knowlton
Access to this links box is available online.