Review

Author for correspondence:
Earl K. Miller
e-mail: ekmiller@mit.edu

Goal-direction and top-down control

Timothy J. Buschman[1] and Earl K. Miller[2]

[1]Princeton Neuroscience Institute and Department of Psychology, Princeton University, Princeton, NJ 08544, USA
[2]The Picower Institute for Learning and Memory and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

TJB, 0000-0003-1298-2761

We review the neural mechanisms that support top-down control of behaviour and suggest that goal-directed behaviour uses two systems that work in concert. A basal ganglia-centred system quickly learns simple, fixed goal-directed behaviours while a prefrontal cortex-centred system gradually learns more complex (abstract or long-term) goal-directed behaviours. Interactions between these two systems allow top-down control mechanisms to learn how to direct behaviour towards a goal but also how to guide behaviour when faced with a novel situation.

1. Introduction

We all have goals—a desired state of the world that we want to achieve. Goals come in many different forms. They can range from short-term, such as finding a snack when hungry, to long-term, such as working towards tenure. Goals can also vary from concrete, such as searching for your keys, to abstract, such as wanting to exercise more. Regardless of their form, all goals share a common thread: one must act on the world in order to achieve them. Therefore, achieving one's goal is by necessity a forward-looking proposition: you try to act upon the world in such a way that its future state will match your desired state.

Such forward-looking behaviour relies on one's previous experience. Through our experiences, we learn how the world behaves and how our actions can change it. The result of this learning is an internalized model of the world, which we can then use to project how the world will behave in the future. For example, if our goal is to find our keys, we might use our previous experience to guide our search towards locations where keys are usually located (e.g. in our bag). Similarly, our internal models can be more abstract, allowing us to generalize them to *guide behaviour in previously unseen circumstances* (e.g. we can find someone else's keys in their office by looking in likely places or following their directions). Application of these models is the essence of 'top-down' or 'executive' control: one must use previous knowledge to plan appropriate actions and then keep 'on task' while achieving the goal. This is the core of intelligent, rational, behaviour—it allows us to not just react to the world but act upon it in order to obtain a desired outcome.

Here, we discuss the neural mechanisms that support such top-down control. Specifically, we suggest goal-directed behaviour uses two complementary systems that can explicitly learn the relationships between actions and outcomes: the basal ganglia (BG) allows simple, fixed, goal-directed behaviours to be quickly learned, whereas the prefrontal cortex (PFC) gradually learns more complex (abstract or long-term) goal-directed behaviours. We then propose a model of how interactions between these two systems can form an 'iterative engine' that, through top-down control mechanisms, directs behaviour towards a goal, even when faced with a novel situation.

2. Top-down control depends on a balance between quick, but concrete, and gradual, but abstract, learning

Goal-directed behaviour relies on learning how to achieve one's goals; that is, learning the relationship between situations, actions and their outcomes. In many situations, one might expect that this learning should proceed as quickly

as possible. Indeed, both humans and animals can quickly learn simple, concrete, associations. For example, a monkey will rapidly learn to associate a particular visual image with a specific, rewarded, motor response (less than 15 trials, e.g. [1]). This certainly makes sense from an evolutionary stand-point: adapting faster than competing organisms will provide a clear advantage in obtaining rewarding outcomes or avoiding harmful ones. However, learning quickly comes at a cost. It may lead to erroneous associations (i.e. coincidences). For example, many of us have experienced superstitious behaviour—an attribution of an outcome to an action despite no mechanistic, causal relationship (e.g. baseball players retightening their batting gloves after every pitch). Similar effects are seen when training artificial neural networks: high learning rates allow the network to quickly converge on a reasonable behaviour, however, (i) learning is often biased towards initial training exemplars and (ii) the network is less stable than networks with lower learning rates (due to the network making big jumps in parameter space with each input; see [2,3] for more on artificial neural networks).

Extending learning over multiple experiences improves the reliability of the association as 'noisy', spurious, correlations are lost and true associations are strengthened. It is also how the 'deep' structure of the world can be discovered. It is the commonalities across experiences that reveal general principles and concepts. Although they come at the cost of requiring more time to develop, such generalized principles have several advantages. First, abstract representations are, by definition, more 'compact' than a more detailed one. Second, generalized principles allow you to act intelligently in a novel situation. Because abstract representations lack non-critical details, they more easily generalize to new circumstances.

3. The neural mechanisms of fast versus gradual learning

Given the advantages (and disadvantages) associated with each form of learning, the brain must balance the pressure to learn as quickly as possible with the benefits of gradual learn-ing. One obvious solution is that both mechanisms are used by the brain, perhaps in complementary neural systems. O'Reilly and co-workers [4] have suggested exactly this type of dynamic exists between fast-learning in the hippocampus and more gradual learning in cortex. Studying the consolidation of long-term memories, McClelland *et al.* [4] suggest fast-plasticity mechanisms within the hippocampus are able to quickly capture new memories while 'training' the slower learning cortical networks. This takes advantage of the special-ization of the hippocampus for rapidly acquiring new information; each learning trial produces large weight changes. They suggest the output of the hippocampus will then repeat-edly activate cortical networks that learn over time with smaller weight changes per episode. Continued hippocam-pal-mediated reactivation of cortical representations allows the cortex to gradually connect these representations with other experiences. That way, the shared structure across experi-ences can be detected and stored, and the memory can be interleaved with others. In the hippocampus, this process pro-ceeds very slowly. For example, bilateral resection of the hippocampus of patient HM resulted in retrograde amnesia for approximately 2 years before surgery, suggesting it takes at least 2 years for full consolidation of memories [5].

This architecture—fast-learning in more primitive, non-cortical structures training the slower, more advanced, cortex—may be a general brain strategy. In addition to being suggested for the relationship between the hippocampus and cortex, it has also been proposed for the cerebellum and cortex [6]. This makes sense: evolutionary pressure on our cortex-less ancestors was presumably towards faster learning, whereas only later were resources available to add a slower, more judi-cious and flexible cortex. We propose that learning of the associations needed for goal-directed behaviour occurs in a simi-lar manner: the BG, a set of subcortical structures, rapidly learn simple associations while gradual learning in the PFC forms more robust, complex and abstract representations.

4. Gradual learning in the prefrontal cortex; fast-learning in the basal ganglia

Goal-directed behaviour relies on the associations learned through previous experiences: we base our estimate of what action is appropriate at the moment on what possible outcome is associated with each possible action. These action-outcome associations were previously captured through learning by strengthening the associations among contexts, actions and out-comes that successfully achieved a goal (i.e. they are rewarded). Conversely, associations that are ineffective at obtaining a reward are weakened. Such learning is 'supervised' in that it requires a teaching signal that reinforces successful associations (and degrades unsuccessful ones). Dopamine (DA) neurons in the midbrain are thought to provide exactly this teaching signal.

Dopaminergic neurons in two midbrain regions (the ven-tral tegmental area, VTA, and the substantia nigra, pars compacta, SNpc), signal a 'reward prediction error' [7,8]. This signal is simply the difference between expected rewards and received rewards. For example, it is positive (and neur-ons are active), when the animal receives an unexpected reward. However, this response disappears if the reward is predicted by a stimulus or an action (as the reward is now expected). Instead, DA neurons will respond to the associated stimulus (or action), as it now 'stands-in' for the reward and occurs unexpectedly [8]. By contrast, if an *expected* reward is *not* received, DA neurons are inhibited, providing feedback that recent behaviour was not effective in obtaining a reward. These reward signals are thought to act as a teaching signal by affecting recently active synaptic connections: strengthening those synapses that are followed by a reward signal while weakening those that do not lead to reward. As we detail next, this simple rule turns out to be incredibly powerful, providing the mechanism that associates context, stimuli and actions that lead to reward.

This teaching signal is thought to act on synapses throughout the brain, but make the largest impact on frontal cortex and the BG since midbrain DA neurons send heavy projections into PFC and the striatum (the main input of the BG, figure 1). Projections into frontal cortex show an anterior to posterior gradient: heaviest in anterior cortex and falling-off as you move posteriorly, suggesting a prefer-ential input of reward information into the PFC relative to posterior cortex [9,10]. Interestingly, the input of midbrain DA into the striatum is much heavier than that of the PFC, by as much as an order of magnitude [11]. Furthermore, DA neurons make connections close to the synapse that stria-tal neurons form with cortical neurons. Indeed, evidence
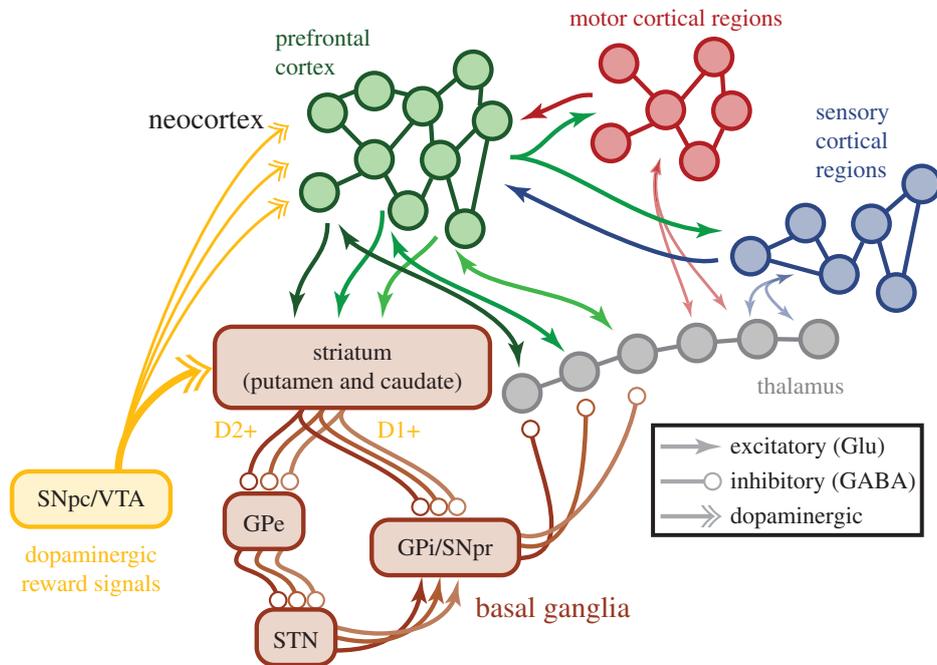
Figure 1. Anatomy of neocortex and BG. Cortical projections enter the BG through the striatum and are thought to maintain separation throughout the BG (shown as different shades). 'Direct' pathway releases inhibition on the thalamus (labelled as D1+); 'indirect' pathway increases inhibition (D2+). Dopaminergic reward signals (in light orange) influence synapses in PFC but are much stronger in striatum. GPe, globus pallidus external parts; GPi, globus pallidus internal parts; SNpr, substantia nigra, pars reticulata; STN, subthalamic nucleus.

suggests that neither strengthening nor weakening of synapses in the striatum by long-term potentiation or depression can occur without DA input [12–14]. By contrast, DA inputs to the cortex are weaker and synapse on the dendrites. Thus, DA may play a strong role in gating plasticity in the striatum while having a more subtle influence in the cortex [15].

We suggest this difference in the way DA influences plasticity in the striatum and PFC leads to a difference in how associations are learned in each region. Specifically, we propose DA strongly influences plasticity in the striatum, producing simple, concrete associations. By contrast, DA has a milder effect in PFC, slowly biasing connections, allowing learning to integrate over many experiences and resulting in a more abstract representation. Differences in learning speed between PFC and striatum were observed in an experiment by Pasupathy & Miller [16]. Monkeys were trained to associate a visual cue with an eye movement either to the left or to the right (in order to receive a reward). Learning occurred over approximately 60 trials, after which the associations were reversed, and the animals had to re-learn new associations. In order to track this learning process across neurons in both PFC and striatum, Pasupathy & Miller recorded from both regions simultaneously. Consistent with a fast-plasticity model of BG, neurons in the striatum showed rapid learning, quickly associating a stimulus with the behavioural response (5–10 trials). By contrast, neurons in PFC showed much slower learning (approx. 30 trials, closely following improvements in behaviour). This difference in learning speed between PFC and BG is consistent with our hypothesis: simple, concrete associations, such as between a stimulus and a motor response, are first identified by the striatum, while slower learning mechanisms in PFC capture associations over a longer time period.

The advantage of such slow-learning in PFC is that it allows learning to integrate over more experiences, constructing a more generalized representation. These generalized representations are crucial to acting appropriately when

faced with a new situation. Evidence for this specific versus generalized trade-off between the striatum and the PFC during learning comes from Antzoulatos & Miller [17], who recorded from multiple electrodes in the lateral PFC and dorsal striatum while animals learned two categories of stimuli. Each day, monkeys learned to associate novel, abstract dot-based categories with a right versus left saccade. Learning occurred over several 'blocks' of trials. During each block, the animal learned the associated movements for a specific set of exemplar stimuli from each category (left versus right). Early blocks of trials consisted of only a few exemplars, allowing the animal to associate each specific *stimulus* with the appropriate response. However, the size of the set of stimuli grew with each additional block during the day, and so the animal was forced to learn the *category* of stimuli in order to make the appropriate response to novel exemplars (i.e. those they had never before seen).

Neurons in PFC and BG played different roles during these two types of behaviours. Early on, when they could acquire specific stimulus–response associations, striatum activity was an earlier predictor of the corresponding saccade. However, as the number of exemplars increased, the monkeys had to form abstractions to classify them. It was at this point that the PFC began predicting the saccade associated with each category (and doing so before the striatum). Thus, it seems that the striatum was leading the acquisition early on when behaviour could be supported by simple stimulus–response learning. However, when the abstraction requirements exceeded that of the simple striatum cache representations, the PFC took over. In this case, the slower learning, associative activity of PFC is ideal for the integration of stimulus properties over many exemplars, allowing for a generalized 'concept' of categories to be learned. This dual-learner strategy allows the animal to perform optimally throughout the task—early on striatum can learn associations quickly while later in the task, when learning associations is no longer viable, PFC guides behaviour.

# 5. Interactions between learning in prefrontal cortex and basal ganglia

So far we have discussed how complementary learning systems in the BG and PFC may allow an animal to balance the need to learn quickly and robustly. However, it is important to note that these systems do not work in isolation. Instead, they are tightly and reciprocally interconnected with one another (figure 1). Indeed, connections between cortex and BG are thought to create segregated 'loops': the connections between nuclei in the BG are maintained such that the output from the BG (via the thalamus) was largely to the same cortical areas that gave rise to the initial inputs into the BG [18]. Owing to the close relationship between cortex and striatum, neurons in both areas tend to have similar responses (e.g. in visual cortex and associated striatum, [19]). Further highlighting the importance of these loops to normal cognitive function, damage to a sub-region of the striatum causes deficits similar to those caused by lesions of its 'looped' cortical region [20,21]. For example, lesions of the regions of the caudate associated with the frontal cortex result in cognitive impairments. In addition, the closely interconnected relationship between BG and PFC would ensure that these regions share information throughout learning.

We should also note that PFC and BG are closely integrated with many other brain regions. For example, the fast-learning system of the hippocampus is known to be closely integrated with PFC and the BG and, like many other brain regions, likely plays a key role in learning and executing behaviours (for review of interactions between hippocampus and BG, see [22]). However, here we focus on the close interactions between PFC and BG and how such a relationship could provide several computational advantages.

6. Learning complex task structures

One outcome of the interaction of different learning styles in PFC and the BG is that they might work together to capture complex task structures. Such an idea was recently put forth in a computational model of task representation by Daw et al. [23]. In their model, Daw et al. represent complex tasks as a decision tree: at each state one can choose between one of many different responses, each of which lead to a new state (with its own stimuli and response alternatives, figure 2). And so, behaviour can be modelled as starting at the topmost 'node' in the tree, choosing a response 'branch', entering a new state, choosing another response, and so on until one has completed the task (hopefully resulting in a reward). Such decision trees underlie many complex behaviours: an initial decision defines what decisions are available in the future. For example, imagine going to work—your initial decision about whether to check the weather before you leave will determine later decisions, like whether you should stop and buy an umbrella when it starts raining.

Daw et al. suggest the flexible associative architecture of PFC is able to capture the entire tree structure, essentially providing the animal with an internal model of the entire task. We propose this representation is due to the combination of (i) how PFC represents information (in high dimensions and over sustained time periods) and (ii) the slower DA-driven learning in PFC (allowing more integrated 'concepts' to be learned). By contrast, Daw et al. suggest BG represents acquired
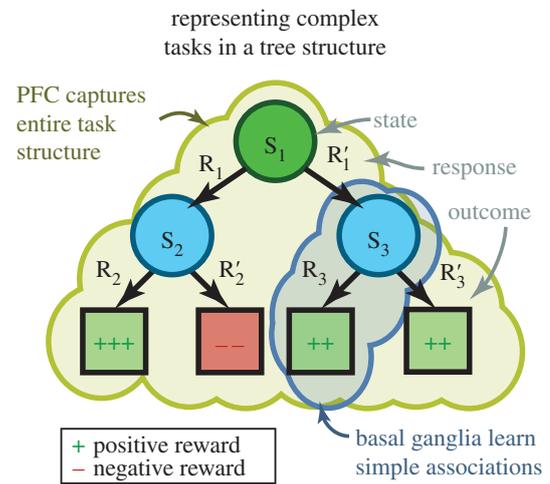

Figure 2. Tree representation of complex tasks. A complex task can be modelled as a set of states (S, circles) with several possible behavioural responses (R, arrows). A response can either lead to a new state or an outcome (squares, either positive, green or negative, red). The fast learning in BG is thought to be ideal for capturing single nodes in the tree (blue cloud) while the slower learning in PFC can capture the entire task (yellow cloud). This more complete view of the task allows for immediately more rewarding responses (e.g. R_1') to be avoided for longer term goals (e.g. R_1 followed by R_2).

information with a 'cache' system that only learns the most rewarding alternative at each decision point (i.e. what is the best branch to take at each state, considered in isolation, without integrating higher order decisions). The BG cache system is computationally simple (and therefore fast) but it is inflexible because the learning is divorced from any change in the outcome. As detailed above, the impact DA has on learning in the BG may be optimized to learn associations in this way.

If true, then this model suggests the initial learning of a complex operant task should begin with the establishment of a simple response immediately proximal to reward (i.e. a single state seen in figure 2). These 'simple' associations are captured by the striatum. Indeed, lesions of the striatum impair learning of new operant behaviours (for review, [24]). More complex tasks, defined as those with more antecedents and qualifications (states and alternatives), involve the PFC to a greater extant. PFC facilitates this learning via its slower plasticity, allowing it to stitch together the relationships between the different states. This is useful because uncertainty of the correct action at any given state adds across the many states within a complex task. Thus, in complex tasks the ability of reinforcement to control behaviour is lessened with the addition of more states. In this situation, the PFC's role may be to build a model of the entire task—complete characterization of how each state relates to another—in order to facilitate learning and guide behaviour (as seen in figure 2). Such models allow for deferral of immediately advantageous states (e.g. S_3 in figure 2) for long-term gain (e.g. taking S_2 to get the maximum reward). In support of this theory, disrupting dorsolateral PFC (via transcranial magnetic stimulation) decreases a subject's ability to use complex models to guide behaviour; instead, subjects tend to choose the immediately advantageous option [25].

As learning progresses, neural representations of tasks are thought to change in one of two ways. Many tasks will remain dependent on the PFC and the models it builds, particularly those requiring flexibility (e.g. when the goal often
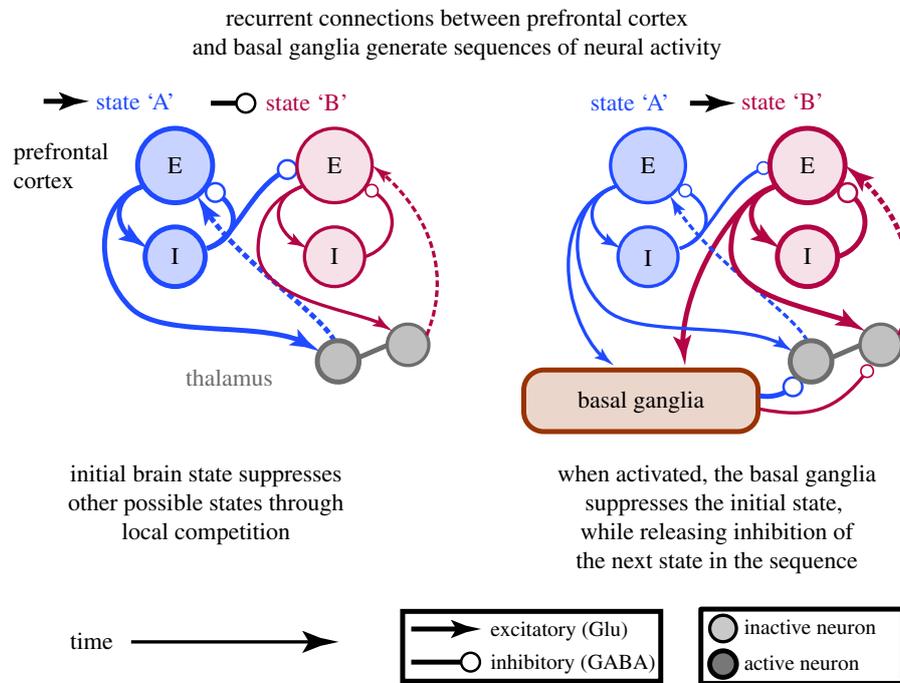
Figure 3. Cortico-ganglia loops support sequencing of behaviours. An initial state in a sequence is cued; activation of this state suppresses alternative states via lateral inhibition. In addition to local processing, the current state is also initially supported via recurrent loops with thalamus (left; dashed lines show ascending projections). The BG gates activity between PFC and the thalamus, acting to inhibit the current state and releasing inhibition on the next state in the sequence (right).

changes) or when the current course is incompatible with a strongly established behaviour (i.e. a habit). However, if a behaviour, even a complex one, is unchanging, then the sequence of appropriate actions can form a new habit. The BG is thought to capture such habits, likely learning them from the patterns of activity in PFC. Indeed, inactivating the BG disrupts *well-learned* motor behaviours, reflecting its continued importance in representing behaviours (note that this continued role for BG is different than the hippocampus, [26]). Furthermore, learning habits *requires interactions* between PFC and BG: optogenetic inactivation of these projections prevented habit formation in a rat trained to run an alternating 'T' maze [27]. However, as noted above, these interactions are not one-way—activity in BG also acts upon PFC. Next, we discuss how such a recurrent connection could be a powerful computational mechanism.

## 7. Computational role of recurrent cortico-ganglia loops

Recurrent connections between cortex and BG also ensure information learned in one region is available in the other. Just as recursion is a powerful algorithmic approach, the recurrent nature of these interactions could serve several different computational purposes in the brain. Here, we briefly outline four possibilities:

First, recurrent connections could allow new experiences to be compared to expectations from previous ones. The resulting 'difference signal' allows for the continued re-evaluation of associations, allowing them to be refined over time. As detailed above, when this difference is between expected rewards and reward outcome, it results in the reward prediction error signal that guides learning. However, the BG may also be activated when other forms of expectations are violated (such as perceptual prediction errors, see [28]). By 'subtracting out'

the expected component of a response, learning can specifically target the un-expected portion without changing the expected (and therefore already understood) components. Similar mechanisms may exist in the cortex [29] and are thought to rely on inhibition in recurrent connections between cortical layers [30]. The anatomical structure and interactions of prefrontal and BG may play a similar role: the inhibitory mechanisms within the BG could act to reduce the activity of expected outcomes.

Second, the loops may allow for sequences of actions (or thoughts) to be strung together (figure 3). Patterns of activity in the PFC descend into BG where associated responses become active. Eventually, this activity leads to the inhibition of the current state (acting via the indirect pathway and propagating to PFC through the thalamus). This, in turn, facilitates the activation of the next associated state in PFC, allowing the animal to prepare for upcoming stimuli, actions, etc. Owing to the recurrent nature of the connections between PFC and BG this process can happen iteratively, allowing a full sequence of actions to be executed in order to achieve a behaviour. Consistent with this view, Barnes *et al.* [31] found that when a rat learns a task, neurons in the striatum become transiently active at each point of the behavioural sequence (e.g. at initiation, when expecting a stimulus, at a decision point and at a reward). This may also underlie the lack of habit formation when PFC−BG interactions are disrupted [27]: without this recurrence the sequences needed for complex habits cannot be generated.

Such iterative sequences may also relate to the oscillatory rhythms observed in PFC and BG (for more on this, see [32]). For example, searching a visual array for a target is an iterative two-step process: you attend to an object and determine whether it is the target. By repeating these steps you will eventually find the searched-for target. Indeed, neurons in a sub-region of PFC, the frontal eye fields (FEF), reflect such an iterative shifting of attention. Interestingly, this iterative
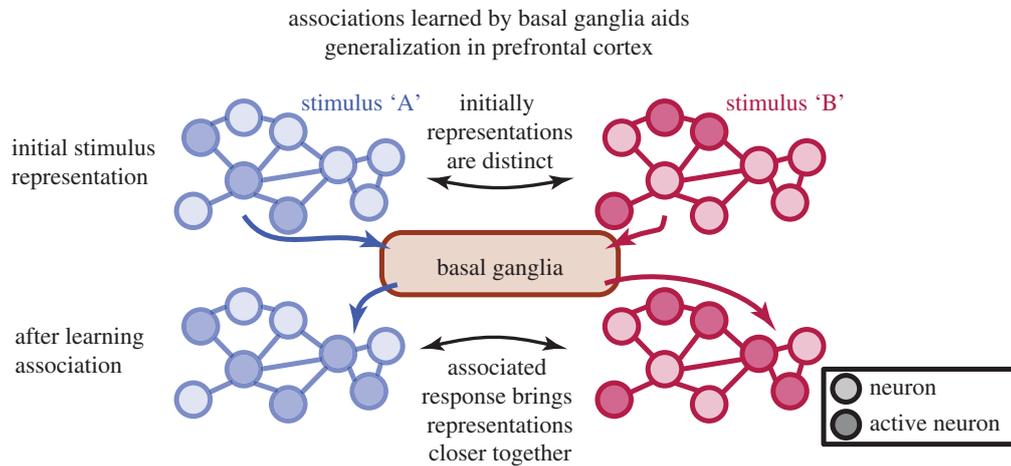
Figure 4. Recurrent learning allows development of complex cognitive representations. Associative learning aids categorization. Initially, stimuli belonging to a common category (e.g. matches and a lighter can both 'start fires') are likely to have disparate representations (top row). However, BG can learn associations between stimuli and behaviours quickly. As this feeds back into PFC, it changes the representation of the stimulus itself, bringing the overall representations closer together (bottom row).
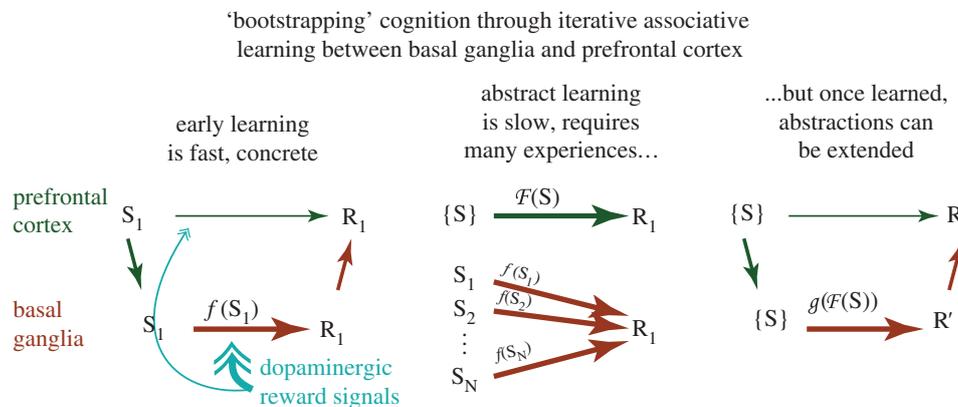
Figure 5. Bootstrapping to developing complex cognitive representations. Recurrence allows for iteratively more complex functions to be learned. Initial learning by BG is fast but concrete (left). Learning is slower in PFC but this allows for more generalized functions to be learned over many different experiences (middle). In this example, a response is learned to a category (set) of stimuli. Once learned the more complex functions/representations are available for further learning by BG, allowing for evermore complex functions to be learned (right).

process was correlated with ongoing beta-band oscillations in PFC, supporting the hypothesis that sequential behaviour may elicit oscillation rhythms in the brain [33].

Third, the recurrent loops between the BG and PFC may support generalization. As noted above, many tasks are learned piecemeal: first singular experiences ('exemplars') are memorized while generalized representations are learned over time. The recurrent connections between PFC and BG may facilitate this process. First, the quick-learning in striatum allows for specific exemplars to be 'mapped' to a behavioural response (figure 4, top). Owing to the recurrent relationship between BG and PFC, the action association learned by BG can become part of the representation in PFC (figure 4, bottom). This association can now act as a 'tag', bringing the representations of multiple exemplars in PFC 'closer' due to their shared association. Once tagged in this way, these patterns will be partially activated during further trials with the same category association, helping learning to shape these representations by finding other commonalities between exemplars.

Finally, the give-and-take in learning between PFC and BG could support model-building itself: once an association is learned in one region it becomes available for further learning in the other. In this way, learning can iterate, allowing for

increasingly complex cognitive representations to be 'bootstrapped' from simpler ones. For example, a well-learned behaviour can form a habit. Habits have the advantage of consolidating representations in the brain, reducing the cognitive load of often-repeated behaviours. In addition, habits can then be used by the executive PFC model-building system as the basis for further learning.

In a similar way, the ability of PFC to learn new categories may be used for richer learning of associations in the BG (figure 5). For example, once the category of 'hammers' is learned, one can instantly generalize associated responses (e.g. hammering a nail) with a new exemplar of a hammer. In other words, the stimulus–response association becomes a more abstract category–response association. Such generalizations are fundamental to cognitive flexibility, allowing one to behave appropriately in new situations. In effect, this process allows for a concept to be elaborated—new actions and new outcomes can become associated with already established categories. However, elaboration is not necessarily limited to a single concept: the recurrent nature of prefrontal and BG interactions may also allow new concepts to be built from already established ones. This 'bootstrapping' process is seen throughout learning and is a hallmark of human

intelligence: we ground new concepts in familiar ones because it eases our understanding of novel ideas. For example, we learn to multiply by serial addition and exponentiation by serial multiplication. Evidence for such 'higher order' representations comes from monkeys trained to perform sequences of movements: while many PFC neurons encoded individual components of a movement sequence, other neurons encoded higher order concepts, such as the type of sequence (e.g. whether it was a sequence of alternating movements or repeated movements, [34]).

8. Using learned associations in prefrontal cortex and basal ganglia to direct behaviour towards a goal

So far we have detailed how associations between stimuli, responses and outcomes can be learned in coordination between PFC and the BG. We then outlined how these complementary systems may explain different aspects of behaviour. Furthermore, we have proposed that recurrent interactions between these systems may provide unique computational advantages, including producing increasingly complex and rich behaviours. However, it is important to note that these representations are learned in service of behaviour: to allow an organism to reliably acquire its goals. Using internalized information such as this to guide behaviour is often referred to as 'top-down' control. Here, we detail the anatomical and physiological characteristics of PFC and BG that allow them to provide such top-down control.

First, both regions are anatomically well-positioned to influence neural activity throughout the brain. For example, PFC is closely integrated with a diverse set of brain regions, sending and receiving projections from most of the cerebral cortex. In addition to interacting with other cortical regions, PFC receives and sends projections to several subcortical regions, including the hippocampus, amygdala, cerebellum, and most importantly for our model, the BG [35–40]. Different PFC subdivisions have distinct patterns of interconnections with other brain systems (e.g. lateral—sensory and motor cortex; orbital—limbic), however, all of these sub-regions are so densely interconnected that any and all information is quickly integrated across sub-regions [41–43]. As noted above, the BG is similarly well-integrated, sending and receiving from most cortical regions. Although these connections do generally form 'loops', connections between nuclei are also partially convergent, possibly providing a mechanism for integration across domains. These diverse connections in both regions may provide the anatomical substrate for top-down control: allowing them to act as a 'hub' of neural processing, synthesizing a wide range of information (sensory, motor, emotional, reward, etc.) and then using this knowledge to influence activity in large swathes of cortex in order to guide behaviour towards a goal.

In addition to being anatomically well-situated, PFC and BG neurons show many of the properties necessary for providing top-down control. First, neurons in both regions sustain their activity across short, multi-second memory delays [44–48]. This 'working memory' property is crucial for goal-directed behaviour, which, unlike 'ballistic' reflexes, typically extends over time and allows associations to be formed between items that are not simultaneously present.

Second, neurons in PFC encode task-relevant information necessary for top-down, goal-directed control of behaviour. This includes task-relevant stimulus information [47], as well as more generalized stimulus information, such as categories [49] or number of stimuli [50]. In addition, as detailed above, PFC neurons encode associations between a stimulus and other stimuli [51] or responses (as detailed above, [1,16]). Furthermore, neurons in orbital PFC will encode reward expectation and uncertainties, signals that are necessary for guiding decision-making in complex tasks (for review, see [52]). The expected outcome following a stimulus and action are also encoded in PFC (with a bias to different sub-regions for each, [53]). Finally, neurons in PFC will encode the current context or situation. Contexts are what capture the contingencies of the situation (i.e. what 'tree' one should be using) and so representing it is necessary to guide behaviour in a goal-direction fashion. Consistent with a primary role in top-down control of behaviour, PFC neurons represent the current context, as well as the 'rules' that govern behaviour in that context [54,55]. Recent work suggests that such contexts are not only represented in single neurons, but in dynamic 'ensembles' of neurons that are defined by synchrony [56]. Such activity-dependent coupling would be highly dynamic, facilitating rapid association between neurons representing stimuli and responses. The speed of these associations, coupled with the ease to re-form new associations, makes synchrony an ideal candidate mechanism of cognitive flexibility.

As detailed above, neurons in PFC and BG are also able to *acquire* such task-relevant information quickly [55,57–59]. This may reflect the fact that PFC neurons are selective for highly complex representations. A recent computational model argued that PFC neurons have 'mixed' selectivity that 'randomly' combine external and internal informations [60]. This results in high-dimensional, sparse, representations throughout PFC. The computational advantage of such a representation is it provides the ability to learn a large number of new associations (perhaps nearly unlimited), including those that were not predicted by the previous structure of the world. Indeed, such mixed selectivity is often observed in PFC (e.g. [51]). Furthermore, coupling such high-dimensionality of representations in PFC [61] with the sustained response of PFC neurons results in an 'evolution' of neural signals in PFC: the current representation of an input depends on previous activity (i.e. the context). This effect was recently highlighted in a working memory task—the response of PFC neurons to an always-irrelevant distractor stimulus depended on the current contents of working memory (even after neural activity levels had returned to baseline, [62]). Such constant integration of context is ideal for top-down control, allowing behaviour to be driven by more than the immediate world.

Together, this brief review highlights how PFC and BG are anatomically and physiologically well-positioned to provide top-down control over neural representations throughout the brain. Direct evidence for PFC's role in top-down control came from a recent study investigating the control of attention. Attention is a commonly studied form of goal-directed behaviour: one attends to a stimulus because it is known to be relevant to the task (and therefore relevant to receiving a reward). Two competing mechanisms are thought to control where/when attention is allocated: attention can be captured in an 'external' way by salient stimuli (i.e. a flashing fire alarm) or can be 'internally' directed towards

task-relevant stimuli. Simultaneous electrophysiological recording of neural activity in parietal and PFC revealed neurons in the FEF sub-region of PFC play a primary role in internally directing attention [63]. Further experiments have confirmed this activity can directly influence neural activity in posterior, sensory cortical regions [64]. These results provide direct evidence that PFC neurons can provide top-down control of neural activity in a goal-directed manner, guided by associations learned in collaboration with BG.

9. Projecting into the future: a proposed neural mechanism for goal-directed behaviour

So far we have outlined how the associations that underlie goal-directed behaviour are learned in complementary systems in PFC and BG. Furthermore, we have provided evidence that these associations could be used to guide behaviour by biasing neural representations throughout the brain. But how can this system be used in a prospective manner? That is, how can we guide behaviour in heretofore unseen circumstances? We propose the same recurrent architecture between PFC and BG that allows for increasingly complex associations to be formed can be run in a 'prospective' way in order to predict the outcome of possible actions. In other words, the current state represented in PFC could be passed through the recurrent loop with BG to generate possible future actions/outcomes. This output can then be captured in the PFC, possibly in parallel with the current state (given the capacity of PFC to simultaneously represent multiple items [65]). Then, if necessary, the entire process can be repeated, allowing for further prospection of outcomes.

This ability probably extends to all recurrent architectures in the brain, allowing previously gained knowledge represented in each sub-system to be used to predict how the world should work in the future. For example, prospection has been seen in the hippocampus: rats will not only think about portions of the environment that they recently experienced but will also project into parts of the environment that are in front of them [66]. Similarly, recurrent connections between PFC and sensory/motor cortex could take advantage of the associations latently learned in these regions. Indeed, asking a subject to 'imagine' a visual scene leads to activation of visual cortex [67]. We propose the same mechanism likely exists between PFC and BG, except this system is designed to predict how current actions predict future outcomes, allowing one to plan a path to one's goal, even when that goal has never been experienced.

References

1. Asaad WF, Rainer G, Miller EK. 1998 Neural activity in the primate prefrontal cortex during associative learning. *Neuron* **21**, 1399–1407. (doi:10.1016/S0896-6273(00)80658-3)

2. Dayan P, Abbott LF. 2005 *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

3. Hertz JA. 1991 *Introduction to the theory of neural computation*. Redwood City, CA: Westview Press.

4. McClelland JL, McNaughton BL, O'Reilly RC. 1995 Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457. (doi:10.1037/0033-295X.102.3.419)

5. Milner B, Corkin S, Teuber H-L. 1968 Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of HM. *Neuropsychologia* **6**, 215–234. (doi:10.1016/0028-3932(68)90021-3)

6. Houk JC, Wise SP. 1995 Feature article: distributed modular architectures linking basal ganglia, cerebellum, cerebral cortex: their role in planning and controlling action. *Cereb. Cortex* **5**, 95–110. (doi:10.1093/cercor/5.2.95)

7. Schultz W. 1997 Dopamine neurons and their role in reward mechanisms. *Curr. Opin. Neurobiol.* **7**, 191–197. (doi:10.1016/S0959-4388(97)80007-4)

8. Schultz W, Apicella P, Ljungberg T. 1993 Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913.

9. Goldman-Rakic PS. 1988 Topography of cognition: parallel distributed networks in primate association cortex. *Annu. Rev. Neurosci.* **11**, 137–156. (doi:10.1146/annurev.ne.11.030188.001033)

10. Thierry AM, Blanc G, Sobel A, Stinus L, Glowinski J. 1973 Dopaminergic terminals in the rat cortex. *Science* **182**, 499–501. (doi:10.1126/science.182.4111.499)

11. Lynd-Balta E, Haber SN. 1994 The organization of midbrain projections to the striatum in the primate: sensorimotor-related striatum versus ventral striatum. *Neuroscience* **59**, 625–640. (doi:10.1016/0306-4522(94)90182-1)

12. Calabresi P, Maj R, Mercuri NB, Bernardi G. 1992 Coactivation of D1 and D2 dopamine receptors is required for long-term synaptic depression in the striatum. *Neurosci. Lett.* **142**, 95–99. (doi:10.1016/0304-3940(92)90628-K)

13. Calabresi P, Pisani A, Centonze D, Bernardi G. 1997 Synaptic plasticity and physiological interactions between dopamine and glutamate in the striatum. *Neurosci. Biobehav. Rev.* **21**, 519–523. (doi:10.1016/S0149-7634(96)00029-2)

14. Kerr JND, Wickens JR. 2001 Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum *in vitro*. *J. Neurophysiol.* **85**, 117–124.

15. Blond O, Crépel F, Otani S. 2002 Long-term potentiation in rat prefrontal slices facilitated by phased application of dopamine. *Eur. J. Pharmacol.* **438**, 115–116. (doi:10.1016/S0014-2999(02)01291-8)

16. Pasupathy A, Miller EK. 2005 Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* **433**, 873–876. (doi:10.1038/nature03287)

17. Antzoulatos EG, Miller EK. 2011 Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron* **71**, 243–249. (doi:10.1016/j.neuron.2011.05.040)

18. Selemon LD, Goldman-Rakic PS. 1985 Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *J. Neurosci.* **5**, 776–794.

19. Brown VJ, Desimone R, Mishkin M. 1995 Responses of cells in the tail of the caudate nucleus during visual discrimination learning. *J. Neurophysiol.* **74**, 1083–1094.

20. Divac I, Rosvold HE, Szwarcbart MK. 1967 Behavioral effects of selective ablation of the caudate nucleus. *J. Comp. Physiol. Psychol.* **63**, 184–190. (doi:10.1037/h0024348)

21. Goldman PS, Rosvold HE. 1972 The effects of selective caudate lesions in infant and juvenile Rhesus monkeys. *Brain Res.* **43**, 53–66. (doi:10.1016/0006-8993(72)90274-0)

22. Johnson A, van der Meer MA, Redish AD. 2007 Integrating hippocampus and striatum in decision-making. *Curr. Opin. Neurobiol.* **17**, 692–697. (doi:10.1016/j.conb.2008.01.003)

23. Daw ND, Niv Y, Dayan P. 2005 Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711. (doi:10.1038/nn1560)

24. Yin HH, Knowlton BJ. 2006 The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* **7**, 464–476. (doi:10.1038/nrn1919)

25. Smittenaar P, FitzGerald THB, Romei V, Wright ND, Dolan RJ. 2013 Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free

control in humans. *Neuron* **80**, 914–919. (doi:10.1016/j.neuron.2013.08.009)

26. Miyachi S, Hikosaka O, Miyashita K, Kárádi Z, Rand MK. 1997 Differential roles of monkey striatum in learning of sequential hand movement. *Exp. Brain Res.* **115**, 1–5. (doi:10.1007/PL00005669)

27. Smith KS, Graybiel AM. 2013 A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron* **79**, 361–374. (doi:10.1016/j.neuron.2013.05.038)

28. Schiffer A-M, Schubotz RI. 2011 Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Front. Hum. Neurosci.* **5**, 38. (doi:10.3389/fnhum.2011.00038)

29. Garrido MI, Kilner JM, Stephan KE, Friston KJ. 2009 The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* **120**, 453–463. (doi:10.1016/j.clinph.2008.11.029)

30. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012 Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711. (doi:10.1016/j.neuron.2012.10.038)

31. Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. 2005 Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* **437**, 1158–1161. (doi:10.1038/nature04053)

32. Buschman TJ, Miller EK. 2010 Shifting the spotlight of attention: evidence for discrete computations in cognition. *Front. Hum. Neurosci.* **4**, 194. (doi:10.3389/fnhum.2010.00194)

33. Buschman TJ, Miller EK. 2009 Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron* **63**, 386–396. (doi:10.1016/j.neuron.2009.06.020)

34. Shima K, Isoda M, Mushiake H, Tanji J. 2007 Categorization of behavioural sequences in the prefrontal cortex. *Nature* **445**, 315–318. (doi:10.1038/nature05470)

35. Amaral DG. 1986 Amygdalohippocampal and amygdalocortical projections in the primate brain. *Adv. Exp. Med. Biol.* **203**, 3–17. (doi:10.1007/978-1-4684-7971-3_1)

36. Amaral DG, Price JL. 1984 Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J. Comp. Neurol.* **230**, 465–496. (doi:10.1002/cne.902300402)

37. Barbas H, De Olmos J. 1990 Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* **300**, 549–571. (doi:10.1002/cne.903000409)

38. Croxson PL *et al*. 2005 Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *J. Neurosci.* **25**, 8854–8866. (doi:10.1523/JNEUROSCI.1311-05.2005)

39. Eblen F, Graybiel AM. 1995 Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J. Neurosci.* **15**, 5999–6013.

40. Porrino LJ, Crane AM, Goldman-Rakic PS. 1981 Direct and indirect pathways from the amygdala to the frontal lobe in rhesus monkeys. *J. Comp. Neurol.* **198**, 121–136. (doi:10.1002/cne.901980111)

41. Barbas H, Pandya DN. 1989 Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* **286**, 353–375. (doi:10.1002/cne.902860306)

42. Pandya DN, Yeterian EH. 1990 Prefrontal cortex in relation to other cortical areas in rhesus monkey: architecture and connections. *Prog. Brain Res.* **85**, 63–94. (doi:10.1016/S0079-6123(08)62676-X)

43. Petrides M, Pandya DN. 1999 Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* **11**, 1011–1036. (doi:10.1046/j.1460-9568.1999.00518.x)

44. Funahashi S, Bruce CJ, Goldman-Rakic PS. 1989 Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349.

45. Fuster JM, Alexander GE. 1971 Neuron activity related to short-term memory. *Science* **173**, 652–654. (doi:10.1126/science.173.3997.652)

46. Levy R, Friedman HR, Davachi L, Goldman-Rakic PS. 1997 Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *J. Neurosci.* **17**, 3870–3882.

47. Miller EK, Erickson CA, Desimone R. 1996 Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* **16**, 5154–5167.

48. Pribram KH, Mishkin M, Enger H, Kaplan SJ. 1952 Effects on delayed-response performance of lesions of dorsolateral and ventromedial frontal cortex of baboons. *J. Comp. Physiol. Psychol.* **45**, 565–575. (doi:10.1037/h0061240)

49. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. 2001 Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312–316. (doi:10.1126/science.291.5502.312)

50. Nieder A, Freedman DJ, Miller EK. 2002 Representation of the quantity of visual items in the primate prefrontal cortex. *Science* **297**, 1708–1711. (doi:10.1126/science.1072493)

51. Rainer G, Rao SC, Miller EK. 1999 Prospective coding for objects in primate prefrontal cortex. *J. Neurosci.* **19**, 5493–5505.

52. Lee D, Seo H, Jung MW. 2012 Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* **35**, 287–308. (doi:10.1146/annurev-neuro-062111-150512)

53. Luk C-H, Wallis JD. 2013 Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. *J. Neurosci.* **33**, 1864–1871. (doi:10.1523/JNEUROSCI.4920-12.2013)

54. Bongard S, Nieder A. 2010 Basic mathematical rules are encoded by primate prefrontal cortex neurons. *Proc. Natl Acad. Sci. USA* **107**, 2277–2282. (doi:10.1073/pnas.0909180107)

55. Wallis JD, Anderson KC, Miller EK. 2001 Single neurons in prefrontal cortex encode abstract rules. *Nature* **411**, 953–956. (doi:10.1038/35082081)

56. Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK. 2012 Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* **76**, 838–846. (doi:10.1016/j.neuron.2012.09.029)

57. Asaad WF, Rainer G, Miller EK. 2000 Task-specific neural activity in the primate prefrontal cortex. *J. Neurophysiol.* **84**, 451–459.

58. Mansouri FA, Matsumoto K, Tanaka K. 2006 Prefrontal cell activities related to monkeys' success and failure in adapting to rule changes in a Wisconsin card sorting test analog. *J. Neurosci.* **26**, 2745–2756. (doi:10.1523/JNEUROSCI.5238-05.2006)

59. White IM, Wise SP. 1999 Rule-dependent neuronal activity in the prefrontal cortex. *Exp. Brain Res.* **126**, 315–335. (doi:10.1007/s002210050740)

60. Rigotti M, Rubin DBD, Wang XJ, Fusi S. 2010 Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front. Comput. Neurosci.* **4**, 24. (doi:10.3389/fncom.2010.00024)

61. Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, Fusi S. 2013 The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590. (doi:10.1038/nature12160)

62. Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J. 2013 Dynamic coding for cognitive control in prefrontal cortex. *Neuron* **78**, 364–375. (doi:10.1016/j.neuron.2013.01.039)

63. Buschman TJ, Miller EK. 2007 Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* **315**, 1860–1862. (doi:10.1126/science.1138071)

64. Moore T, Armstrong KM. 2003 Selective gating of visual signals by microstimulation of frontal cortex. *Nature* **421**, 370–373. (doi:10.1038/nature01341)

65. Buschman TJ, Siegel M, Roy JE, Miller EK. 2011 Neural substrates of cognitive capacity limitations. *Proc. Natl Acad. Sci. USA* **108**, 11252–11255. (doi:10.1073/pnas.1104666108)

66. Davidson TJ, Kloosterman F, Wilson MA. 2009 Hippocampal replay of extended experience. *Neuron* **63**, 497–507. (doi:10.1016/j.neuron.2009.07.027)

67. Kosslyn SM, Thompson WL, Klm IJ, Alpert NM. 1995 Topographical representations of mental images in primary visual cortex. *Nature* **378**, 496–498. (doi:10.1038/378496a0)