The role of prediction and outcomes in adaptive cognitive control

Anne-Marike Schiffer [*1], Florian Waszak [2,3] and Nick Yeung[1]

[1]Department of Experimental Psychology, University of Oxford, OX1 3UD, UK
[2]Université Paris Descartes, Sorbonne Paris Cité, Paris, France
[3]Laboratoire Psychologie de la Perception, UMR 8158, Paris, France

Conflict of Interest:
The authors declare no competing financial interests.

Corresponding author:
[*] anne-marike.schiff er@psy.ox.ac.uk

Abstract

Humans adaptively perform actions to achieve their goals. This flexible behaviour requires two core abilities: the ability to anticipate the outcomes of candidate actions and the ability to select and implement actions in a goal-directed manner. The ability to predict outcomes has been extensively researched in reinforcement learning paradigms, but this work has often focused on simple actions that are not embedded in hierarchical and sequential structures that are characteristic of goal-directed human behaviour. On the other hand, the ability to select actions in accordance with high-level task goals, particularly in the presence of alternative responses and salient distractors, has been widely researched in cognitive control paradigms. Cognitive control research, however, has often paid less attention to the role of action outcomes. The present review attempts to bridge these accounts by proposing an outcome-guided mechanism for selection of extended actions. Our proposal builds on constructs from the hierarchical reinforcement learning literature, which emphasises the concept of reaching and evaluating informative states, i.e., states that constitute subgoals in complex actions. We develop an account of the neural mechanisms that allow outcome-guided action selection to be achieved in a network that relies on projections from cortical areas to the basal ganglia and back-projections from the basal ganglia to the cortex. These cortico-basal ganglia-thalamo-cortical 'loops' allow convergence - and thus integration - of information from non-adjacent cortical areas (for example between sensory and motor representations). This integration is essential in action sequences, for which achieving an anticipated sensory state signals the successful completion of an action. We further describe how projection pathways within the basal ganglia allow selection between representations, which may pertain to movements, actions, or extended action plans. The model lastly envisages a role for hierarchical projections from the striatum to dopaminergic midbrain areas that enable more rostral frontal areas to bias the selection of inputs from more posterior frontal areas via their respective representations in the basal ganglia.

Keywords:

Cognitive Control; Reinforcement Learning; Hierarchical Reinforcement Learning; Action Selection; Prediction; Ideomotor Principle; Basal Ganglia; Striatum; Dopamine; Prefrontal Cortex

Adaptive human behaviour relies on the capacity to select and perform actions in accordance with desired outcomes. This requires at least two abilities: the ability to predict outcomes, and the ability to select actions (or sets of actions) on the basis of these predictions to achieve specific goals. These two aspects have to some degree been researched separately in neuroscience and psychology. The ability to predict outcomes has been extensively investigated using learning paradigms (O'Doherty, 2004; Pessiglione et al., 2006; Tricomi and Fiez, 2008; Walsh and Anderson, 2012) in both human subjects (Holroyd and Coles, 2002; O'Doherty, 2004) and animals (Matsumoto et al., 2009; Schultz and Dickinson, 2000; Wise, 2004). Here, the main focus has been the prediction of primary rewards or aversive events (Hikosaka et al., 2008; Matsumoto and Hikosaka, 2007; Schultz and Dickinson, 2000) and the mechanisms of reinforcement learning that translate these predictions into observed behaviour at the level of individual stimuli and specific actions (Schultz, Dayan, and Montague, 1997). This line of work has paid much less attention to the question of how action selection is guided by high-level goals and by outcomes that inform on the successful completion of an action in the absence of measurable reward.

The ability to guide behaviour in accordance with high-level goals has received substantial scrutiny in human cognitive neuroscience. Here the focus has been on the ability to select and implement *task sets* that specify appropriate mappings from environmental stimuli to behavioural responses to suit current task demands (Monsell, 1996), particularly in the presence of salient distractors or habitual alternate actions (Miller and Cohen, 2001; Norman and Shallice, 1986), subsumed under the concept of *cognitive control* (Monsell and Driver, 2000). However, this framework has tended to neglect a fundamental feature of behaviour, that actions are usually directed toward achieving particular outcomes. In contrast, many commonly used cognitive control paradigms—such as task switching, response conflict, and response inhibition—feature responses that have no direct consequences. Instead, execution of the required response is considered to represent completion of the task.

Here we explore recent ideas that promise to bring together prior research on outcome prediction and cognitive control and thus contribute to the development of unifying accounts of adaptive action selection. In the first half of this article, we briefly review foundational concepts in research on cognitive control and reinforcement learning, before presenting hierarchical models of reinforcement learning as a promising framework for linking ideas from these two hitherto rather separate domains of research. A key feature of these hierarchical models is their proposal that actions are not selected in isolation, but instead are learnt and selected in structured sequences that are directed toward identified goals. Crucially, these structured sequences bear strong conceptual

resemblance to the notion of task sets in theories of cognitive control. However, an important limitation in current theories of hierarchical reinforcement learning is that they fail to specify clearly the neural mechanisms supporting sequential action selection. In the second half of this article, we propose that structured sequencing of actions depends critically on the basal ganglia and their interactions with frontal cortex. We review neurophysiological, anatomical, and neuropsychological evidence in support of this claim. We conclude by outlining promising directions for future research to test these proposals.

1 Cognitive Control

1.1 Goal-directed behaviour

In research on cognitive control, a founding observation is that human behaviour is highly flexible and, as a consequence, highly under-constrained by the environment (Allport, 1980; Miller and Cohen, 2001; Monsell, 1996; Norman and Shallice, 1986). Presented with stimuli as simple as written words on a computer screen, for example, experimental subjects are capable of an enormous variety of responses: reading the words aloud, counting their syllables, counting the vowels, providing rhymes, judging whether they refer to concrete objects or abstract ideas, giving synonyms, searching for particular target letters, etc., etc. Given this flexibility, mechanisms of cognitive control are required to guide action selection according to current goals and intentions (Miller and Cohen, 2001; Norman and Shallice, 1986), a function of particular importance when a number of possible actions have to be coordinated. This coordination is thought to depend on establishing an effective organization—a *task set*—that specifies the stimuli to be attended, the type of response to be made, and the appropriate mapping between these stimuli and responses (Monsell, 1996).

A classical test of cognitive control is therefore the task-switching paradigm (Jersild, 1927), in which participants are presented with a series of stimuli (such as digits) and are required to make rapid and flexible switches between two or more tasks (such as judging whether the digit is odd or even, versus judging whether it is greater or less than 5). Subjects are typically told which task to perform for each presented stimulus, either trial-by-trial through instructional cues or according to a predictable schedule, but are sometimes given freedom of choice. Regardless, each task switch requires them to disregard the current task set and establish a new one, a process associated with a cost in performance and associated activity across a network of frontoparietal

cortical regions implicated in cognitive control (Dosenbach et al., 2006; Duncan, 2010; Kim et al., 2012; Richter and Yeung, 2014). The established task set is needed to facilitate repeated performance of the same tasks on multiple stimuli, and to shield performance against interference from competing stimulus-response mappings (Collins and Frank, 2013; Collins, Cavanagh, and Frank, 2014; Dreisbach and Haider, 2008).

Many other well-established cognitive control paradigms share this focus on 'willful' goal-driven task implementation, including the go/no-go, Stroop, and Eriksen flanker tasks, to name a few. These tasks all share the critical feature that one of a number of possible responses must be selected or withheld (go/no-go) according to a cue (task-switching, go/no-go) or predefined rule (Stroop, flanker), often in the face of competition from more habitual responses to that stimulus (Stroop). Evidence from these various tasks has converged on a core set of processing principles: first, that behaviour emerges from competitive interactions among representations of potential stimuli and responses; second, that this competition occurs simultaneously at multiple levels of processing (e.g., in the visual system from representations of simple features to complex objects); and, finally, that competition operates under the guiding influence of goal representations in prefrontal cortex (Desimone and Duncan, 1995). This research has been highly successful in characterising the computational (Miller and Cohen, 2001) and neural (Sakai, 2008) bases of this top-down influence. Specifically, it is commonly held that prefrontal cortex maintains stable representations of task-relevant information and current task goals (a working memory function). This information is held to modulate processing in sensory and motor cortices, as well as the interaction between them, in service of effective task performance.

1.2 Hierarchical structure

Accumulating evidence suggests that cognitive control is not a unitary construct but can instead be fractionated into interacting component processes with distinct and identifiable neural bases. For example, interactions between prefrontal cortex and the medial temporal lobe appear to support retrieval of goals, set during earlier prospective planning (Cohen and O'Reilly, 1996; Schacter, Addis, and Buckner, 2007), with input from orbitofrontal cortex providing information about the likely payoffs of those plans (Koechlin and Hyafil, 2007) and input from medial prefrontal cortex providing more negatively valenced information about costs and uncertainty (Behrens et al., 2007; Botvinick, 2007).

Of particular relevance to the present discussion is the proposal that regions in prefrontal cortex are specialised for representing actions at different levels of abstraction (Badre and

D'Esposito, 2009; Koechlin and Summerfield, 2007; but see Duncan, 2010 for a dissenting view). It has long been recognised that behaviour is hierarchically structured, with high-level plans (e.g., get to work on time) comprised of sub-routines (e.g., make breakfast, pack a bag, walk to the office) that themselves involve lower-level sub-routines (e.g., make coffee, toast bread). These abstraction hierarchies nearly always exhibit a sequential structure, such that sub-routines are executed in a natural or necessary ordering at each level of the hierarchy. As such, action selection can be characterised in terms of activation flow in a hierarchy from the high-level plan down to the particular sequences of concrete actions that are ultimately specified for execution (Lashley, 1951; Miller, Galanter, and Pribram, 1960).

This form of hierarchical structure is common in cognitive theories of action selection (Botvinick and Plaut, 2004; Cooper and Shallice, 2000). Recent neuroimaging evidence suggests that corresponding hierarchical structure is explicitly represented along the rostro-caudal axis of lateral prefrontal cortex, with more rostral regions containing increasingly higher-level representations of actions. For example, whereas activity in premotor cortex is observed when coloured cues indicate the required response, activity in the inferior frontal gyrus becomes apparent when coloured cues indicate the overall task (i.e., how to respond to other stimulus attributes) rather than a particular action, while activity in rostral prefrontal cortex only becomes apparent when these cue-task contingencies switch (Koechlin, Ody, and Kouneiher, 2003). Collectively, this evidence suggests an important elaboration of the concept of cognitive control, in which goals are simultaneously represented at multiple levels of abstraction in distinct parts of prefrontal cortex, with influence flowing down hierarchically from high-level plans to specific implemented actions.

1.3 Concrete goals and sequential actions

The work sketched in the preceding subsections is illustrative of substantial progress made in our understanding of the computational and neural mechanisms of cognitive control. However, these successes notwithstanding, a striking feature of many standard cognitive control paradigms is their narrow focus on tasks in which actions are produced in a stimulus-driven and reactive manner, with success in the task defined in terms of producing pre-defined responses to given stimuli rather than in terms of bringing about a desired state of affairs in the world (Hommel, 2009). For example, in prior research on flexible task switching, the moment of response execution is typically taken to be the end of the trial. Even trial-to-trial feedback is rarely provided, a choice that follows the lead of three studies that laid the methodological foundations of this work

(Allport, Styles, and Hsieh, 1994; Meiran, 1996; Rogers and Monsell, 1995 - only Meiran's Experiment 4 included trial-to-trial feedback of any kind). Given this, it is perhaps not surprising that few studies since have explicitly manipulated action outcomes in task switching, with the notable exception of experiments using adaptations of the Wisconsin Card Sorting Task in which subjects must use trial-to-trial feedback to infer a sorting rule (e.g., Monchi et al., 2001; Rogers et al., 1998) and a handful of studies using reward incentives to motivate effective switching (Kleinsorge and Rinkenauer, 2012; Nieuwenhuis and Monsell, 2002; Shen and Chun, 2011). References to feedback and action outcomes are notable by their absence in recent authoritative reviews of task-switching research (Grange & Houghton, 2014; Kiesel et al., 2010; Vandierendonck, Liefooghe, and Verbruggen, 2010) as well as in the related research literatures on response conflict (MacLeod, 1991; Yeung, 2013) and response inhibition (Aron, Robbins, and Poldrack, 2014). One domain in which the role of outcomes in cognitive control has been studied in more detail is in terms of performance feedback (Ridderinkhof et al., 2004; Walsh and Anderson, 2012), but with a few important exceptions (Ribas-Fernandes et al., 2011; Krigolson and Holroyd, 2006; Collins and Frank, 2013), this work has focused on learning and optimisation of simple stimulus-response associations rather than high-level task sets. Moreover, feedback in these studies is typically used to indicate whether the subject produced the required response to the imperative stimulus, rather than being a meaningful consequence of the particular action produced.

In stark contrast, in almost all complex everyday behaviours, our actions are instrumentally directed towards achieving certain desired outcomes or producing specific changes in the environment. Thus, successful completion of an action is typically defined in terms of bringing about its desired outcome, not in terms of its execution per se (i.e., we define success in terms of ends not means). Success is a light turning on rather than a switch being flicked, a hot cup of coffee in our hands rather than completion of pouring and stirring actions, a draft manuscript rather than a long sequence of keypresses. Action outcomes of this sort have no obvious correlate in many cognitive control tasks, for which action execution marks the end of the trial and in which feedback is often not provided (and, when provided, may be unnecessary except during the earliest stages of practice; Holroyd and Coles, 2002). As such, previous research on cognitive control may have neglected a critical route by which tasks and actions are selected, that is, through the outcomes they are intended to achieve.

A second key feature of everyday action that is missing from standard cognitive control paradigms is sequential structure. With a few notable exceptions (e.g., Schneider and Logan,

2006), these paradigms involve a series of discrete trials in which no explicit structure governs the relationship between successive events. Indeed, the most common approach is to deliberately randomise the order of presented stimuli and required tasks and responses (Richter and Yeung, 2014). This design choice is true even for paradigms that notionally tap hierarchical behavioural structure (e.g., Koechlin, Ody, and Kouneiher, 2003). As such, these paradigms share little in common with everyday behaviour which, as already noted above, is characterised by hierarchical and sequential structure: Plans at a given level of abstraction typically comprise a series of sub-routines for which the order is at least somewhat constrained (e.g., in my morning routine, I must get out of bed before I can make coffee or shower, but the order of the latter two sub-routines can be exchanged; Botvinick and Plaut, 2004). Action outcomes play a critical role in this behavioural sequencing, because it is often the case that those outcomes are necessary preconditions for later actions in a sequence: success in picking up a spoon allows me next to add coffee grounds to the pot.

Thus, while cognitive neuroscience research has been highly successful in characterising the computational mechanisms and neural basis of control, it has achieved this in the context of task paradigms that neglect crucial features of human behaviour. This narrow focus may at least partly explain why lesion studies have often found little or no substantive impact on cognitive control of damage to regions that consistently show control-related activation in imaging studies (Holroyd and Yeung, 2012). This is not to say, however, that there has been no research on the outcome-driven, sequential nature of behaviour; indeed this has been a major focus of research, albeit largely separate from the work reviewed above. It is to this research that we now turn.

2 Action Outcomes

2.1 Action effects

Actions typically have sensory consequences: some that are intrinsic to the intended goal (e.g., a light turning on at the flick of a switch), some that are intrinsic to the action itself (e.g., the proprioceptive consequences of finger flexion), and some that are incidental but nevertheless consistently associated (e.g., the auditory click of the switch). A large corpus of findings has documented the formation of associations between actions and these sensory consequences (Herwig and Waszak, 2009; Herwig, Prinz, and Waszak, 2007; Kühn et al., 2010; Waszak et al., 2005; Krieghoff et al., 2011), for example as investigated according to the ideomotor principle of

action selection (James, 1890; Prinz, 1990; Greenwald, 1970; Hommel, 2009). According to this research, sensory consequences of actions become part of the internal representation of actions (Hommel, 2009; Band et al., 2009) and therefore play a crucial role in action selection (e.g. Hommel, 2009). For example, after learning that particular actions have consistent sensory consequences (e.g., pressing a key produces a low-frequency tone), presentation of those 'consequences' prior to action selection biases action selection toward the associated action (Elsner and Hommel, 2001).

The acquisition and maintenance of functional representations necessitate not only the ability to predict which action leads to which outcome, but also the ability to assess whether events and outcomes concur with original predictions. This mechanism of prediction and evaluation is reminiscent of forward models in motor control (Blakemore and Sirigu, 2003; Wolpert and Miall, 1996; Wolpert and Kawato, 1998). In these models, every mismatch between predicted and actual sensory feedback is indicative of an error and hence the need for adjustment. These forward models can also usefully be run offline, without actual movement. Through this mechanism, they can be used to estimate what the sensory consequence of a possible movement would be, and thereby select and adjust movements based on the predicted divergence between this anticipated consequence and the planned motor command (Blakemore and Sirigu, 2003; Miall, 2003; Schaal, Mohajerian, and Ijspeert, 2007; Wolpert and Miall, 1996; Wolpert and Kawato, 1998). There is ample evidence for neural activity corresponding to such outcome-centred routines of prediction, evaluation, and adjustment in motor control (Tunik, Houk, and Grafton, 2009).

However, this research has to date made limited contact with studies of cognitive control. In the present context, the critical missing conceptual link is the idea that action-effect predictions ought to depend on the overarching task goal in two key respects. First, as noted above, some sensory consequences are intrinsic to the action whereas others are incidental, and we might expect differential processing of these even if the action-effect correlations are equivalent. There is some evidence on this point (e.g., Krigolson and Holroyd, 2006), but little systematic study. Second, predictions ought to be task-dependent. For example, the same action (e.g., a flick of switch) will have different predicted consequences in the context of different tasks (e.g., turning on a light vs. turning on an electrical socket), yet action-effect bindings are commonly studied in the context of tasks with fixed action-effect associations. As such, while studies of sensory action effects clearly demonstrate the principle that outcomes play a critical role in adaptive action

selection, this research has less to say about the role of action outcomes in the higher-level control of behaviour.

2.2 Reward and reinforcement learning

Perhaps the best-characterised form of action-outcome learning relates to associations involving motivationally salient events: reward and punishment. The topic of reinforcement learning has become a major focus in neuroscience research, with interest fuelled by the striking convergence between formal computational theory and observed properties of neural systems underpinning learning in humans and other animals. Reinforcement learning theory (RL; Schultz, Dayan, and Montague, 1997; Sutton, 1988; Sutton and Barto, 1990; Sutton and Barto, 1981) explains how agents (humans/animals) learn to choose actions that will maximise their future rewards (Barto and Simsek, 2005). The agent assigns a value to a given state, which signifies how much reward this state predicts, and learns which actions lead to transitions from one state to the next. In accordance with Thorndike's *law of effect* (Thorndike, 1927), actions that lead to rewarding states are more likely to be repeated in the future.

Take for example a case of searching for groceries in a foreign country on a Sunday, when shops are closed. Your foraging may eventually lead you to a gas station, which you enter, and wherein to your surprise you find food and drink. This experience may change the value you ascribe to gas stations, and make trips to the gas station more likely on future Sundays. RL describes learning processes such as this: If reward is delivered unexpectedly, the present state will be assigned a higher value than it had before (because it is now recognised as a state that may yield reward). At the same time, the value of the state(s) that preceded the present state will also be increased because they predict future reward. Actions that lead to states that promise reward will be produced more often. The unexpected delivery of reward is called a positive reward prediction error, because a state yielded more reward than previously expected. The next time the same sequence of states is encountered, reward delivery will have been predicted to some degree (depending on the learning rate) by the previous state. This makes reward less unexpected and hence decreases the positive reward prediction error (Schultz et al., 1992; Schultz, Dayan, and Montague, 1997; Schultz, 2007). To refer back to our example, on another foraging trip a week later, finding food at the gas station would not be such a large positive surprise, but finding the gas station to be closed might prompt disappointment (negative prediction error).

The aim in RL is to choose the action that will lead to the highest expectation of future reward (discounted for time-to-reward). The underlying routine can be understood as a constant

prediction of future outcomes paired with the evaluation of the present state with regards to whether it yielded the outcomes it was predicted to yield. Deviations from predictions cause prediction errors and result in adjustment of the model, i.e., learning.

2.3 Hierarchical reinforcement learning

Hierarchical reinforcement learning (HRL) is a development of RL theory that aims to deliver computationally more tractable solutions for complex environments than 'flat' RL. Crucially for our purposes here, HRL theory also has interesting implications for the role of outcome and task representations in cognitive control (Collins and Frank, 2013**; Collins & Koechlin, 2012)**. One computational issue for reinforcement learning is that its sequential nature poses a scaling problem (see Botvinick, Niv, and Barto, 2009, for review). The computational demands for the learning agent increase with each step and each action, rendering complicated multistep actions computationally unfeasible. HRL circumvents this problem by enabling actions to be selected and learnt in coherent sequences. For example, within the options framework (Botvinick, Niv, and Barto, 2009; Holroyd and Yeung, 2012), primitive actions can be compiled into higher-level mini policies, or *options*, that guide the selection of action sequences. Because the action sequence encompassed in each option is treated as a chunk, it can be selected in its entirety, as opposed to selecting single actions at a time, with the potential to greatly simplify learning in complex task domains. Indeed, options can themselves be sequenced and compiled into higher-level options, allowing for yet further simplification of the task space. Importantly, learning occurs according to the same principles as standard RL: options that produce better-than-expected outcomes are reinforced, while successful completion of a selected option reinforces preceding lower-level actions. In this way, learning is achieved simultaneously at multiple levels of abstraction, identifying high-level options as well as low-level actions that produce positive outcomes.

Options are associated with probable end-states, which can be conceived of as subgoals. Subgoals are anticipated outcomes, but not necessarily primary rewards: They can be states that allow the selection of other actions that will ultimately deliver reward. This feature is of great relevance. As discussed previously, we choose actions to achieve desired outcomes, which in everyday life may not necessarily be primary rewards. If we start the day by choosing to go into work to teach a class, arriving at work is an important subgoal that informs us we are on the right track. Arriving at work, however, is not itself a primary reward. Attaining or failing to attain a subgoal is informative and relies on the same routine of prediction and evaluation as appraisal of

primary reinforcers does in RL (Diuk et al., 2013; Ribas-Fernandes et al., 2011). The idea here is that a wide array of sensory inputs can fulfil a function similar to that of reward, if these inputs deliver reliable information about the state of the environment (or the agent in the environment). Meanwhile, many human studies use abstract feedback symbols such as point scores, colours, or icons which may (Holroyd et al., 2004; Nieuwenhuis et al., 2005; Yeung and Sanfey, 2004), but need not (Elliott, Frith, and Dolan, 1997; Klein et al., 2007; Nieuwenhuis et al., 2005; Swanson and Tricomi, 2014; Ullsperger and von Cramon, 2003), represent a monetary outcome (see Kringelbach et al., 2003; Seitz, Kim, and Watanabe, 2009 for paradigms using actual primary reward). Clearly, these symbolic sensory events need to be cognitively interpreted as valuable states to assume the characteristics of reward. The degree of abstraction and the temporal extension of the plans humans pursue, such as succeeding as a teacher or publishing an article, call for neural mechanisms of learning that can signal whether events occur as expected, even if they are temporally and/or conceptually distant from primary reward delivery. In sum, a crucial feature of HRL is that it encompasses the concept of reaching informative states, which permit evaluation of a completed set of actions and afford the selection of a new series of actions to approach an overarching goal. This concept is intuitively appealing when considering the role of outcomes in everyday actions. The HRL framework also promotes the idea that neural structures known to code for reward prediction may be involved more generally in event/state prediction and coding of mismatches between anticipated (intended) states and actual outcomes.

2.4 HRL and cognitive control

Research on RL and its recent extension in HRL has been exceptionally successful in characterising the acquisition and evaluation of reward predictions. Crucial for our purposes, HRL is also of direct relevance to addressing the limitations of cognitive control research identified above. There, we noted that this research has adopted a narrow conception of ‘goals’ that does not capture the intuitive notion that goals fundamentally relate to states of the world brought about by our actions, and that it relatedly fails to capture the idea that human behaviour is intrinsically structured and sequential in nature. HRL addresses precisely these features: Within RL, actions are reinforced in proportion to their individual propensity to bring about states of the world associated with positive outcomes; within HRL, positive reinforcement can occur for structured sets and sequences of actions, and can be brought about in terms of informative as well as rewarding world states. Thus, actions in HRL are truly goal-directed. They are also structured, by virtue of being compiled into nested hierarchies of options.

Importantly, the concept of options bears close resemblance to that of task sets (Collins and Frank, 2013) studied in research on cognitive control, particularly in light of recent proposals about hierarchical representations in prefrontal cortex (Botvinick, Niv, and Barto, 2009; Collins, Cavanagh, and Frank, 2014). This parallel has been developed elsewhere to link computational theories and neuroimaging studies of hierarchical control (Botvinick, 2008), and to reconcile apparent discrepancies between imaging and lesion data on the neural mechanisms of cognitive control (Holroyd and Yeung, 2012). These proposals have emphasised key similarities between options and task sets: Both are representations that are abstracted over low-level actions to specify coherent groupings of responses; both are proposed to be selected and activated in a hierarchical fashion, with flow-down of activation that guides selection of increasingly concrete and specific action plans; and both are held to guide behaviour over extended periods.

There is clear convergence in terms of the proposed neural underpinnings of HRL in recent accounts (Botvinick, Niv, and Barto, 2009; Holroyd and Yeung, 2012). In particular, high-level option representations are proposed to depend crucially on lateral prefrontal cortex (Botvinick, Niv, and Barto, 2009), with input from orbitofrontal and perhaps medial prefrontal regions (Holroyd and Yeung, 2012), and with representations in these regions influencing action selection in the basal ganglia under the influence of reward. This network description has recently been extended to include the proposal that the orbitofrontal cortex (OFC) represents states that afford specific options (Stalnaker et al., 2014; Wilson et al., 2014). In this way, knowledge of the physiology of prediction, evaluation, and action selection from reinforcement learning research may deliver valuable insights into the mechanisms by which outcomes may globally influence action selection and cognitive control. However, extant proposals have had little to say about the neural mechanisms underpinning a fundamental component of the HRL system: the sequencing of actions within a chosen option. This is the question addressed in the second half of this article, in which we propose a model of the neuroanatomy and neurophysiology of sequential structure in adaptive action selection.

3 The neurophysiology of outcome predictions

3.1 Dopaminergic signalling of prediction error

The role of dopamine and dopaminergic projection pathways in RL was established by the seminal finding that positive reward prediction errors lead to phasic increases in cell firing in the ventral

tegmental area and substantia nigra in the midbrain (Schultz et al., 1992; Schultz, 2000; Schultz, Dayan, and Montague, 1997; Suri, 2002). These midbrain structures are the primary source of dopamine in the brain (Bjoerklund, 2007; Haber, 2003 for review). One of the major projection pathways of the dopaminergic system is the mesostriatal pathway (Bjoerklund and Dunnett, 2007 for review), which targets the striatum, a nucleus in the basal ganglia (Bédard et al., 1969; Haber, 2003). Similar to the nigral and tegmental dopaminergic cell assemblies, their projections to the striatum have repeatedly been shown to be involved in the coding of prediction errors (Daw et al., 2011; Joel, Niv, and Ruppin, 2002; O'Doherty, 2004; Schultz and Dickinson, 2000). Intact mesostriatal projections seem to be pivotal in learning from feedback (Holl et al., 2012; Shohamy et al., 2008). More recently, it has been shown that prediction errors in hierarchical reinforcement settings, which concern predictions of the value of options, are computed in the striatum (Daw et al., 2011; Diuk et al., 2013; Jin, Tecuapetla, and Costa, 2014). These findings have fostered the view that the striatum plays a special role in selecting actions and evaluating their outcomes (e.g., Frank, Scheres, and Sherman, 2007; Houk et al., 2007; Redgrave, Prescott, and Gurney, 1999).

If hierarchical prediction error coding is important for outcome prediction in actions, one would expect to find signatures of prediction errors when actions do not yield desired or anticipated outcomes, that is, when subgoals are not achieved. Importantly, this neural signature should be present even if these subgoals are not directly related to primary reward (cf. Torrecillos et al., 2014). It is not yet conclusively established that striatal prediction errors are observed when subjects fail to reach anticipated subgoals (end states of options) that never entail reward delivery, and do not change the overall estimate of reward likelihood. However, a few studies investigating prediction errors in perception have yielded evidence that the striatum codes for the unexpectedness of events *per se* (den Ouden et al., 2009; Grahn, Parkinson, and Owen, 2008; Grahn and Rowe, 2013; Schiffer and Schubotz, 2011; Schiffer et al., 2012; Seger et al., 2013) and is not limited to reward-related prediction error coding. Although unexpected events in these studies were not predictive of forthcoming reward, or positive feedback, they were sometimes task-relevant (e.g., Schiffer and Schubotz, 2011), even if only to the degree that they informed participants that they should pay attention to deviations in a stimulus to increase their ability to answer (unrewarded) questions correctly (Schiffer et al., 2012). This finding stands in contrast to the idea that striatal prediction errors code solely for changes in the expected sum of future rewards. Rather, it favours the idea that the striatum codes in a model-based fashion for the (un-) expectedness of events more broadly (Schultz, 2013). One rationale behind this claim is that humans need to learn about unexpected deviations from their expectations to adapt behaviourally

to lasting changes in the environment (Behrens et al., 2007; Schiffer et al., 2013). The degree to which these events need to be related to future reward probability in a complex model of the world requires further testing.

Continuing interest in the computations of the dopaminergic midbrain and striatum is also fuelled by the finding that dopamine's action on a number of sites within the basal ganglia appears crucial in establishing associations between cortical representations (such as motor commands and sensory outcomes) and in choosing actions based on these representations (Bischoff-Grethe, Crowley, and Arbib, 2002; Frank, 2005; Gurney, Humphries, and Redgrave, 2015; Redgrave, Prescott, and Gurney, 1999; Stocco, Lebiere, and Anderson, 2010). To understand how basal ganglia anatomy may hence contribute to goal-directed action selection, we next consider three prominent features of its neurophysiology:

1. Cortical projections ('loops') through the basal ganglia enable associations between activation patterns in different cortical areas.
2. Projection pathways within the basal ganglia play a role in the acquisition and selection of extended sequences.
3. Interactions of hierarchical projections in the basal ganglia through subcortical loops allow context-dependent modulation of task sets.

3.2 Prediction in cortico-basal ganglia-thalamo-cortical loops

We propose that cognitive control in sequential goal-directed actions is subserved by cortico-basal ganglia-thalamo-cortical loops, a system proposed to compute predictions of sensory states online and to link representations of actions to their sensory correlates (Bischoff-Grethe, Crowley, and Arbib, 2002; Gurney, Prescott, and Redgrave, 2001; Redgrave, Prescott, and Gurney, 1999). The basal ganglia comprise 9 subcortical nuclei: the caudate nucleus, putamen, and nucleus accumbens (N.Acc), which are together referred to as the striatum, and the globus pallidus externa (GPe), globus pallidus interna (GPi), subthalamic nucleus (STN), substantia nigra pars reticulata (SNr), substantia nigra pars compacta (SNc), and ventral tegmental area (VTA) (Figure 1; Graybiel, 1998; Saint-Cyr, 2003).

A first criterion for a neural structure that can function as a sequential predictive action control system is the capacity to integrate input from a wide range of cortical areas, for example from sensory, motor, and multimodal association cortices. In fact, the striatum as the input structure to the basal ganglia shows such a remarkable pattern of connectivity. Virtually the entire

neocortex projects to the striatum. It thus receives projections from sensory cortices (e.g., Seger, 2008), as well as motor and premotor areas and prefrontal sites (Di Martino et al., 2008; Kemp and Powell, 1970; Selemon and Goldman-Rakic, 1985; Parent and Hazrati, 1995; Schmahmann and Pandya, 2008; Schmahmann and Pandya, 2006). For some of these areas, a very distinctive kind of projection pathway has been established in terms of cortico-basal ganglia-thalamo-cortical loops (Alexander, DeLong, and Strick, 1986; Haber, 2003; Parent and Hazrati, 1995; Selemon and Goldman-Rakic, 1985). A key characteristic of these loops is that cortical input areas project to specific, circumscribed areas within the striatum. The striatal area sends even more converged projections to the output nuclei of the striatum, the GPi and SNr (Figure 1A). The information is then transferred via the thalamus back to one of the cortical input regions (Alexander, DeLong, and Strick, 1986; but see Joel and Weiner, 2000). For example, the motor loop, as first described in the monkey (Alexander, DeLong, and Strick, 1986), has inputs from the supplementary motor area, the arcuate premotor area, the motor cortex, and the somatosensory cortex. These projections converge in the same area of the putamen. The putamen then projects to the ventrolateral GPi and caudolateral SNr. The projection from these output nuclei reaches two specific thalamic nuclei. Lastly, the thalamo-cortical projections of the motor loop terminate in the mesial premotor cortex (supplementary motor area). The same principle can be found in all cortico-basal ganglia-thalamo-cortical loops (Alexander et al., 1986). However, input areas are not necessarily adjacent areas of neocortex: the executive loop receives input from the dorsolateral prefrontal cortex (dlPFC), posterior parietal cortex (Figure 1A), and arcuate premotor area, all of which target the same area in the dorsolateral head of the caudate nucleus, which in turn projects via the thalamus back to the dlPFC (Alexander, DeLong, and Strick, 1986; Selemon and Goldman-Rakic, 1985). Projections from non-adjacent areas of the cortex into overlapping or interdigitating areas of the striatum are one aspect of basal ganglia neuroanatomy that may contribute to the role of this system in integration of information.

The proposed predictive control system incorporates the characteristic of sequentiality connecting predictions of present states, motor intentions, and sensory outcomes. Thus, a second criterion for a system supporting predictive sequential control is access to representations of planned movements, as well as their end states, which in turn form the preconditions of subsequent actions within a sequence. One important aspect of the loop structure satisfying this condition is that output states are fed back into the system (Berns and Sejnowski, 1998; Stocco, Lebiere, and Anderson, 2010). On the motor level, for example, projections of the representation of the hand and digits in the motor cortex (handknob; Yousry et al., 1997) interdigitate in the

striatum with the projections from the hand area in primary sensory cortex (S1; Graybiel, 1998; Flaherty and Graybiel, 1991). Projections from different areas reach the dendritic spines of so-called medium spiny neurons within the same patches in the striatum (matrisomes, cf. Graybiel, 1998; Flaherty and Graybiel, 1991; Flaherty and Graybiel, 1993). This physiological property of projections to the striatum may provide neural templates (Graybiel, 1998) for the association of different cortical input and output patterns (Houk and Wise, 1995; Graybiel, 1998), in a process modulated by dopamine (see below).

On a more general level, these forward models can be described as the association of a motor command, action, or choice (Houk and Wise, 1995; Houk et al., 2007) with sensory states (Bischoff-Grethe, Crowley, and Arbib, 2002) or multimodal representations (cf. Stocco, Lebiere, and Anderson, 2010). In cognitive terms, we can thus rephrase these associations as iteratively linking the representation of a present state (or context, Apicella, 2007) with sequential action possibilities to reach a desired goal state. Sequential representations in the basal ganglia are known as chunks (Graybiel, 2008; 2005; 1998a). This functional description and terminology dovetails with the concepts of chunks in HRL, denoting an action sequence that can be treated as an entity.

A third condition for a system that selects sequential (chunked) action is that predictions of action outcomes need temporal precision, for example because delayed sensory input may be indicative of failed actions (cf. Sardo et al., 2000). Importantly, associations between converging inputs in the striatum may be linked within a defined time window (provided in the striatum by so called tonically active neurons, *TANs*; Morris et al., 2004; Sardo et al., 2000). It has been proposed that activation in the striatal projection neurons is modulated by activity levels of TANs (Reynolds, Hyland, and Wickens, 2001; Reynolds and Wickens, 2002; Apicella, 2007). Specifically, cessation of TAN activity may allow striatal projection neurons to become active and transmit information. TANs may thus provide a timestamp for associations to become active and acquired (Smith et al., 2004; see Stocco, Lebiere, and Anderson, 2010 for a computational implementation), allowing predictive sequential models of motor command copy, anticipated sensory consequence, and subsequent motor command copy, wherein states are kept separate to allow successive implementation. It is important to keep in mind that while this example focuses on the motor loop, the same principle holds for loops originating in prefrontal areas, with sensory state representations that are likely to be multimodal (Saint-Cyr, 2003; Seger, 2008).

A fourth criterion for a control system of sequential actions is the ability to detect deviations from intended sequences and signal these deviations to allow behavioural adjustments.

An emerging view is that the basal ganglia are involved in selecting action sequences (Graybiel, 1998) and monitoring to detect deviations both within the sequence, as well as at the outcome level (Carr, 2000; Graybiel, 1998; Grahn and Rowe, 2013). As we have seen, basal ganglia connectivity provides the essential features to support the monitoring of sequentiality (Stocco, Lebiere, and Anderson, 2010), based on its ability to associate sequential representations from various cortical inputs while discriminating serial positions. Mismatch signals in the basal ganglia may hence code for deviations from the sequence even if they do not change the predictions of future primary reward. The dopaminergic signal established in (H)RL may be one example of such a signal of deviation. We next describe why dopaminergic prediction errors in the basal ganglia may specifically play an essential role in extended sequential action, owing to dopamine's role in the acquisition and selection of weighted forward models.

3.3 Probabilistic selection in basal ganglia pathways

The ability to select appropriate actions to achieve internal goals is fundamental to cognitive control. Because most contexts are preconditions for a number of actions, action selection must take into account how likely it is that each possible action is appropriate (Collins and Koechlin, 2012; Donoso, Collins and Koechlin, 2014) and will yield the anticipated outcome. Action selection can be efficient and fast if different alternatives are associated with weights that encompass this probability, and selection occurs based on these weights. Within the basal ganglia, acquisition of these probabilistically weighted forward models is held to depend on two key organisational features: the distribution of the dopaminergic receptors on medium spiny neurons in dual projection pathways in basal ganglia loops, and the consequences of dopaminergic action on these receptors.

Within each cortico-basal ganglia-thalamo-cortical loop, there are three separate projection pathways, of which two have opposing effects on cortical activity. These two pathways are associated with different dopaminergic receptors. One type of dopamine receptor (D1 type), located on dendrites of medium spiny neurons, gives rise to the direct projection pathway (Albin, Young, and Penney, 1989; Smith et al., 1998; Bolam et al., 2009). This projection pathway reaches the output structures GPi and SNr monosynaptically, which in turn project to the thalamus. Activation of D1 receptors disinhibits the thalamus via this direct projection pathway. Disinhibition of the thalamus increases cortical activity, such that activation in the direct pathway leads to increased activity of corresponding cortical output patterns. This pathway has thus been dubbed the 'go pathway' (Figure 1A; Frank, 2005; Frank, Seeberger, and O'Reilly, 2004).

Conversely, striatal neurons that express another type of dopamine receptor, so-called D2 receptors, send projections to the GPe. The GPe projects to the STN, which in turn projects to the GPi and SNr output nuclei and thence to the thalamus. If dopamine binds to D2 receptors, these indirect pathway projections lead to inhibition of thalamic activity and accordingly no increase in cortical activity. Therefore, the indirect projection pathway via the GPe and STN has been called the 'no go pathway' (Frank, Seeberger, and O'Reilly, 2004; Figure 1b). The third pathway, the hyperdirect pathway, will not be discussed further, but its relevance to action selection has been described elsewhere (e.g., in Frank, 2006; Jahfari et al., 2012; Nambu, Tokuno, and Takada, 2002; Nambu, 2004). Importantly, current models suggest that each forward model is represented simultaneously in separate sets of medium spiny neurons within the direct and indirect pathways (Frank, 2005; Gurney, Humphries, and Redgrave, 2015). This means that activation of each cortical representation depends on the dominance of its respective representation in the 'go' pathway compared to the 'no go' pathway.

Historically, models of the direct and indirect pathway have focussed on their role in selecting actions in relation to current dopamine levels. However, recent models that very successfully predict behaviour in patients with Parkinson's disease (PD; Frank, Seeberger, and O'Reilly, 2004; Frank, 2005; Frank, 2006; Frank et al., 2007) have focused on the role of dopamine bursts in shaping learning in the basal ganglia, thus influencing the probabilistic dominance of 'go' and 'no go' activity in future action selection (Frank, 2005; see Gurney, Humphries, and Redgrave, 2015 for a very recent model which compares striatal action selection and learning). Specifically, activation of D1 receptors on the medium spiny neurons, which give rise to the direct pathway, results in long-term potentiation (LTP) of synaptic efficacy. At the same time, D2 receptor activation prevents LTP on these synapses (Reynolds and Wickens, 2002). If a representation of cortical activity in the striatum is accompanied by a dopamine burst, this hence leads to increased synaptic strength for the representation in the direct pathway and possibly a concurrent decrease in synaptic strength for the representation in the indirect pathway. Thus, in the motor domain, dopamine bursts teach both pathways to make one response more likely, while concurrent alternative responses are suppressed (Morris, Schmidt, and Bergman, 2010; Frank, 2005 for a review). With reference to the RL literature, these dopaminergically modulated synaptic weight changes can explain why rewarded actions are chosen with increasing probability (Bogacz and Gurney, 2007; Botvinick, Niv, and Barto, 2009; Frank and Claus, 2006; Gurney, Prescott, and Redgrave, 2001; Gurney, Humphries, and Redgrave, 2015).

LTP in the two-pathway account is very useful in explaining how forward models can be strengthened. However, perhaps due to a historical view of the basal ganglia as a pure motor structure, and perhaps because dopaminergic modulation has been predominantly associated with reward (but see Horvitz, 2000 for an early dissenting view; Redgrave and Gurney, 2006 for a review), action selection in the basal ganglia pathway has often been taken to be the exact opposite of cognitive control: Selection in the basal ganglia has been associated with reward-oriented habitual responses, whereas cognitive control focuses on the implementation of novel or instructed tasks, especially if the task is not directed at the most salient stimuli or most dominant response. However, dopamine-driven action selection can be reconciled with the proposed role of the basal ganglia circuits in outcome prediction for non-rewarding action outcomes if the hierarchical nature of actions and projection pathways in the basal ganglia is taken into account (Graybiel, 1998).

As we will see, recent descriptions of basal ganglia connectivity (Haber, 2003; Draganski et al., 2008) point to an involvement in cognitive control that is even more intriguing than monitoring of sequential steps and acquisition of probabilistic forward models for rewarded actions. Specifically, these structures have been implicated in mediating top-down control by anterior prefrontal areas (representations of abstract, high-level goals) over more posterior areas in frontal cortex (representations of concrete actions). These recent findings suggest interplay between projections from input areas holding representations at different levels of action hierarchies (Haber, 2003; Koechlin and Hyafil, 2007). As we will discuss next, understanding these hierarchically organised striato-nigral loops may contribute substantially to our understanding how information flows from areas representing abstract action plans to areas representing lower-level actions (Badre and D'Esposito, 2009; Koechlin and Hyafil, 2007).

3.4 Multiple projection hierarchies in the basal ganglia

Recent physiological and computational investigations of basal ganglia connectivity have focussed on interconnectivity between the striatum and the SN/VTA complex (Haber, 2003; Draganski et al., 2008; Haruno and Kawato, 2006). These subcortical loops provide a compelling explanation of the impact of orbitofrontal and prefrontal inputs on biasing action selection in the striatum (Frank and Claus, 2006; Desrochers and Badre, 2012; Haber, 2003; Haruno and Kawato, 2006; Karamati & Gutkin, 2013). A simplified account of the role of these subcortical loops is that the ventromedial striatum influences the dorsolateral striatum via its projections to the dorsal tier of the substantia nigra, which sends dopaminergic projections to the dorsolateral striatum.

Dopaminergic innervation of the dorsolateral striatum is thus under the influence of the ventromedial striatum. This is relevant because different parts of the prefrontal cortex project to the ventromedial and dorsolateral striatum (Alexander, DeLong, and Strick, 1986; Crittenden and Graybiel, 2011; Desrochers and Badre, 2012; Draganski et al., 2008; Eblen and Graybiel, 1995; Haber, 2003; Figure 1B).

A very detailed model of these projections (Haber, 2003), recently supported by diffusion tensor imaging (DTI; Draganski et al., 2008), describes these striato-nigral loops as spiralling downwards from striatal projection zones corresponding to anterior prefrontal areas to projection zones corresponding to posterior prefrontal and motor areas. This spiral largely follows the rostro-caudal axis of the PFC that has been related to the level of abstraction of representations (Badre & D'Esposito, 2009; Koechlin, Ody, and Kouneiher, 2003). Because the OFC and ventromedial prefrontal cortex (vmPFC) lie anterior to the dlPFC, and their projection zones in the striatum follow a similar gradient, the described projections enable striatal projection zones of the OFC and vmPFC to modulate dopaminergic innervation of the dlPFC's striatal projection zones via the striato-nigral loops. The striatal projection zone of the dlPFC in turn modulates dopaminergic projections to the striatal projection zone of premotor cortex. Because dopamine is relevant to the selection of representations, these projections may provide a mechanism by which rostral (higher-order) prefrontal representations bias selection in more caudal areas (Figure 1B).

Within our HRL framework, modulation of striatal representations of dlPFC input by OFC projections is particularly interesting in relation to the concepts of option availability and of action selection within options (Stalnaker et al., 2014; Wilson et al., 2014). The OFC has been associated with coding the present states, i.e., states affording different options (Wilson et al., 2014). Support for this claim comes from a study showing that OFC codes for transitions between states that afford different options, even if each state has the same reward value (Stalnaker et al., 2014). In this study, Stalnaker and colleagues used single-cell recordings to show that OFC neurons do not only convey information about reward value. Rather, they also signalled the beginning of a new experimental block, even if reward value remained unchanged while sensory features of the rewards changed.

In contrast to this role for OFC, the dlPFC has been associated with representation of rules that guide actions (Wilson et al., 2014). One tantalising idea is that the OFC representation of the current state can bias selection of actions represented in the dlPFC via its connectivity to the striatum and midbrain dopaminergic system, initiating weight-changes between different probabilistic forward models in the striatal pathways (Badre and Frank, 2012; Frank and Badre,

2012). A recently presented model by Frank and Badre (2012) exploits the hierarchical setup of fronto-striatal and striato-nigral loops to explain hierarchical action selection in a reward context. Their model posits that more rostral frontal areas influence which striatal representations of input from more caudal cortical areas are facilitated in the direct pathway. In support of this model, the same authors presented fMRI data in a companion paper (Badre and Frank, 2012) suggesting that contextual representations of (hidden) task rules in the prefrontal cortex bias the selection of premotor representations in the basal ganglia.

Both lines of research (Badre and Frank, 2012; Frank and Badre, 2012; Stalnaker et al., 2014) suggest that multiple subcortical projection pathways are involved in action selection of nested hierarchies (spanning from states signifying available options, via representations of currently active task rules, to motor commands). A similar hierarchical account of cortico-subcortical projection pathways has recently been used to explain drug-seeking behaviour (Keramati & Gutkin, 2013), which was previously discussed in terms of reinforcement learning (Wise, 2004). These findings warrant further research, not least because the role that specific frontal areas play in this rostro-caudal axis of the fronto-striatal loop remain a matter of debate: Very similar functions have been ascribed to different cortical areas (e.g., in Badre and Frank, 2012 and Wilson et al., 2014) and, conversely, dissimilar functions have been proposed for nearly identical areas of cortex (compare for example Badre and Frank, 2012 and Derrfuss et al., 2005). Further, while some studies point towards involvement of the basal ganglia in sensory predictions and sensory prediction-error coding (den Ouden et al., 2009; Grahn et al., 2013; Schiffer and Schubotz, 2011; Schiffer et al., 2012), it is yet to be tested empirically whether neural networks involved in HRL support outcome prediction and action selection in non-reward contexts as proposed. Each of these questions of functional neuroanatomy needs to be followed up in future research.

4 Outcomes in Cognitive Control

In first three sections of this review we have presented evidence for the role of outcomes in action selection, and their relative neglect in cognitive control paradigms. We have discussed how the role of outcomes in action selection is well-established in RL and HRL. Moreover, by considering the apparent relevance of basal ganglia projection pathways in reward prediction, we have described how the physiological properties of the basal ganglia could support the implementation and evaluation of chunked sequential actions. The remainder of this review will discuss the

proposition that cognitive control—the ability to implement necessary actions to achieve desired outcomes—should benefit from explicit representation of outcomes. We will review evidence for the importance of an intact basal ganglia network for cognitive control and outcome anticipation, then discuss additional evidence for the role of outcomes in control based on behavioural and electrophysiological paradigms in healthy subjects. Lastly, we will consider the characteristics of experiments that might uncover the putative benefits of outcome representations in adaptive cognitive control.

4.1 Evidence for basal ganglia contributions to cognitive control and outcome prediction

Recent computational models stress the role of the basal ganglia in cognitive control (Gurney, Prescott, and Redgrave, 2001; Hazy, Frank, and O'Reilly, 2007; O'Reilly and Frank, 2006; O'Reilly, 2006; Stocco, Lebiere, and Anderson, 2010). Evidence from research into the cognitive changes experienced by Huntington's disease (HD) and Parkinson's disease (PD) patients support these models. Both of these neurological conditions affect the basal ganglia and lead to cognitive as well as motor impairments. Whereas HD is signified by cell death in the striatum (Kowall, Ferrante and Martin, 1987), the primary neural signature in PD is a loss of dopaminergic cells in the substantia nigra (Bernheimer et al., 1973; Riederer and Wuketich, 1976).

Perhaps the most frequently studied cognitive control functions in these patients are task switching and response inhibition in go/no-go paradigms. Both HD and PD patients show larger behavioural costs of task switching (Aron et al., 2003; Cools, 2006; Cools et al., 2001; Cools et al., 2003; Holl et al., 2012; Shook et al., 2005). PD patients' task-switching deficit appears particularly pronounced for switches to non-habitual behaviour compared to habitual behaviour (Cameron et al., 2010). This selective impairment supports the argument that the projections of the basal ganglia's dopaminergic system are involved in selecting appropriate actions, rather than habitual responses. Deficits in cognitive control paradigms such as the Stroop colour-word task, the trailmaking test, and the Tower of London task are also particularly well-established for Parkinson's disease (Nobili et al., 2010; Robbins et al., 1994). The Tower of London task requires participants to plan multiple steps ahead before implementing a movement sequence. The fact that PD patients take longer for this planning phase (Robbins et al., 1994), during which subsequent states and movements must be emulated and evaluated, is particularly interesting given the idea that the basal ganglia play an important role in outcome anticipation. Moreover, PD patients do not show predictive strategies in motor tasks (Crawford et al., 1989; Flowers, 1978). Lastly, error

detection is compromised in PD (Ito and Kitagawa, 2006) and in patients with focal basal ganglia lesions (Ullsperger and von Cramon, 2006), similarly indicating a compromised ability to evaluate action outcomes.

Results implicating the basal ganglia in outcome prediction have also been obtained by Holl and colleagues (2012). The authors showed that the presence of feedback in a probabilistic classification learning task (the Weather Prediction Task; Knowlton, Squire, and Gluck, 1994) determines whether the basal ganglia will be recruited. They also replicated an earlier finding that patients with basal ganglia impairments are particularly impaired in implicit learning from feedback (Shohamy et al., 2008 for review). Another intriguing result for the role of the basal ganglia in task switching comes from a TMS study (van Schouwenburg et al., 2012) showing that dopamine levels in the putamen influence task-switching abilities. Collectively, these studies suggest that compromised performance in cognitive control tasks associated with basal ganglia disorders may be explained by the structure's role in outcome prediction and evaluation. In addition to this clinical evidence, we will now discuss empirical evidence for the relevance of outcomes on performance from cognitive control paradigms directly.

4.2 Evidence for the role of outcomes in action selection

Our hypothesis is that representations of predicted outcomes should play a key role in selecting extended sequential actions. While this key question has rarely been addressed in research on cognitive control to date, the role of outcomes for selection of non-sequential, individual actions has been investigated within the framework of ideomotor control theory (here, outcomes are often referred to as action effects). As reviewed above, the ideomotor principle posits that anticipated action effects (sensory consequences of actions) are incorporated into the representation of actions, creating a bi-directional link between actions and action effects (Band et al., 2009; Herwig, Prinz, and Waszak, 2007; Hommel, 2009; James, 1890; Prinz, 1990; see Lukas, Philipp, and Koch, 2012; Janczyk, Heinemann, and Pfister, 2012; Gaschler and Nattkemper, 2012; Ziessler, Nattkemper, and Vogt, 2012 for recent applications of the paradigm).

Solid evidence has been gathered that the predictability of a sensory consequence limits the surprise response to these effects (Band et al., 2009; Bednark et al., 2013), in line with the proposal that outcome representations are activated when actions are selected. Band and colleagues (2009) implemented auditory sensory outcomes in a four-response task-switching paradigm. They did not find that responses with predictable outcomes were performed faster, but did observe response slowing on trials following unexpected outcomes. Neural recordings

dovetailed with the behavioural results, in that unpredictable outcomes elicited an event related potential (ERP) component resembling those seen following rare negative feedback. This finding was recently substantiated by a study showing that unexpected action outcomes elicit larger ERPs when they carry informative value concerning the correctness of a response (Bednark et al., 2013). Although these studies emphasise the association between actions and outcomes, they typically do not report a benefit of outcome predictability at the selection stage (Band et al., 2009; Lukas, Philipp, and Koch, 2012). On the other hand, as reviewed below, a few recent studies have shown that anticipating an outcome can facilitate action selection (Marien, Aarts, and Custers, 2012; Ruge, Müller, and Braver, 2010; Ziessler, Nattkemper, and Vogt, 2012).

Together, these existing fragments of evidence begin to suggest that anticipation or presentation of action outcomes can influence high-level action selection. Adding arbitrary but predictable sensory effects to task sets may have little effect on participants' ability to perform the task but may delay performance of the following response. However, it cannot be ruled out that this effect is owing to increased processing demands, or a reorienting effect caused by the surprise. In fact, the described cortico-basal ganglia-thalamo-cortical loop model of action selection suggests that ordinary sensory effects of button presses should suffice to inform participants about the correctness of their response as long as the rules of the task set are known. According to the model, for each response set, representations of the stimuli are associated with representations of their respective correct responses (motor command and expected sensory feedback) within the striatum. Failure of the correct sensory feedback to occur, or an unpredicted sensory event occurring, would elicit a prediction error that can be used to evaluate the correctness of the response (cf. Holroyd and Coles, 2002).

4.3 Evidence that meaningful outcomes facilitate associated actions

The studies reviewed in the previous section found reliable but limited impact of arbitrary sensory action effects. In contrast, more robust outcome effects have been observed in a set of studies for which outcomes are inherent in (or intrinsic to) the task and are meaningfully related to the actions performed (Marien, Aarts, and Custers, 2012; Ruge, Müller, and Braver, 2010; Ziessler, Nattkemper, and Vogt, 2012). Methodologically, the studies are very different from each other, but they share the critical common feature of having semantic coherence between actions and outcomes. It seems plausible that this core feature of everyday behaviour is essential to bringing about effects of outcome utility in future cognitive control paradigms.

One study showing that selection of simple actions benefits from the presence of contingent outcomes was reported by Ziessler and colleagues (2012). This study found a behavioural benefit of a match between an imperative stimulus and an action effect, using object and grip-type affordances as ‘go’ stimuli and sensory effects in a go/no-go paradigm (Ziessler, Nattkemper, and Vogt, 2012). When images of objects were used as stimuli, actions were performed faster if pictures of object-compatible grip types served as the action effect. The same effect was established when grip types were used as ‘go’ stimuli and compatible objects as action effects. This paradigm thereby exploits overlearned associations between objects and grip types. Facilitation of action execution by presentation of objects has previously been shown (Grèzes et al., 2003), but the study is interesting in that the authors show that this priming effect can be used to elicit an arbitrarily associated response. While this type of finding is usually taken to corroborate the ideomotor principle (Ziessler, Nattkemper, and Vogt, 2012; cf. Kunde, 2001), we argue that anticipation of outcomes facilitates action selection (as posited by ideomotor theory) and further allows the agent to chunk actions into nested hierarchies and to monitor their successful completion (as implied by the application of HRL principles to cognitive control).

Facilitation of action selection at higher levels of organization has been observed in an fMRI study by Ruge, Müller, and Braver (2010), who showed that the cost of switching tasks is slightly but reliably decreased when feedback for a response (sensory outcome) is delivered in the same (spatial) dimension as the original response, compared to feedback with no spatial or response-related properties. Switches in the spatial feedback condition activated a neural network associated with allocation of spatial attention in a personal reference frame comprised of anterior intraparietal sulcus, dorsal premotor cortex and rostral cingulate zone (Ruge, Müller, and Braver, 2010).

Moreover, a study by Marien, Aarts and Custers (2012) investigated the effect of changing the colour used to identify the target stimulus in a pair of letters. The crucial manipulation was whether subjects represented the task in terms of its goal—to classify the letter as a vowel or consonant—or in terms of its means—to focus on a particular colour (by virtue of the instructions they were given). Goal representations led to more effective switching, particularly when pre-switch responses were rewarded (Marien, Aarts, and Custers, 2013; 2012).

Within the framework we present, these effects can be explained by the fact that in all studies participants were motivated to chunk representations of motor commands together with representations of indicators of successful actions beyond the re-afference delivered by the motor response. Ziessler et al. (2012) achieved this by exploiting learned associations between stimuli

that act as preconditions to specific responses and said response. Ruge and colleagues suggest that their participants experienced spatially compatible visual outcomes of their responses as effects of their actions. Marien, Aarts and Custers (2012) prompted participants to use representations of higher-level action-outcomes for action selection, using instructions that emphasized the goal of the task as opposed to the stimulus-response mapping. These interpretations are in line with a computational account which suggests that action selection requires basal ganglia-driven updating of working memory representations (Frank, Loughry, and O'Reilly, 2001; O'Reilly, 2006; O'Reilly and Frank, 2006; Donoso, Collins, and Koechlin, 2014). Translated to the terminology of this model, the present claim would be that chunking of stimulus-response mappings into options (task sets) yields performance benefits in all three paradigms because working memory updating of options is not required in switch trials when outcomes are represented at a higher level of abstraction, but updating is required when switching between stimulus-response mappings. To return to our initial example: on return visits to different countries we do not need to retrace our steps along specific routes to buy dinner; instead, we are able to acquire, select, and switch fluently between different situation-specific options for achieving a particular overarching goal, such as choosing to buy food at a gas station in one country versus a convenience store in another.

5 Conclusion and a look ahead

We have outlined a hypothesis that situates basal ganglia function within an HRL framework to integrate existing work on action-effect binding, reinforcement learning, and cognitive control. In the present review we have explored the relationship between these concepts, to develop the proposal that prediction and evaluation of outcomes—specifically, of subgoals within action sequences—underpin the selection and monitoring of extended sequential actions. In line with recent proposals about the computational and neural basis of HRL, we propose that extended action sequences can be chunked and then represented and selected as coherent options. On the basis of computational models and anatomical properties of the basal ganglia, we argue that control of chunked, extended sequential actions relies on dopaminergic modulation of a network of recurrent loops connecting cortical and subcortical components, with this control extending beyond situations where subgoals are signified by primary reward.

This proposal has two key implications. The first is in terms of identifying a template for future research on the topic of outcome-guided sequential action. Classical paradigms may not be able to show the benefit of prediction on the action level, because sensory consequences of responses already hold sufficient information to evaluate task performance. New paradigms to

investigate the role of outcomes in cognitive control must contain features that allow testing of the following behavioural predictions:

- Introducing predictable outcomes into cognitive control tasks should improve action-selection compared to absent or unpredictable outcomes, but only if those outcomes provide performance feedback that is not encompassed fully in the sensory feedback of the action.
- In task switching designs, transitions to new tasks should be performed faster if information about task-outcome contingencies is available than if the relationship between tasks and outcomes is ambiguous.
- In a sequence of tasks, progression from one task to the next should be faster if the end of each subtask is signified by an individual outcome, as opposed to identical outcomes for different subtasks.
- In a sequence of tasks, confusion of subtasks (errors of order) should be decreased if each subtask is signified by an individual outcome, as opposed to identical outcomes for different subtasks.
- Patients suffering from neurological disorders of the basal ganglia should show reduced modulation of task performance by introducing predictable outcomes than healthy controls in tasks with established outcome effects on cognitive control.

The other key implication of the proposed model is the requirement to develop novel paradigms to substantiate the proposed basal ganglia-mediated interactions between rostral prefrontal/orbitofrontal and more posterior prefrontal sites in extended action selection, and to determine the exact role of specific subcortical projection pathways in top-down biased action selection. Of particular interest is the investigation of the most anterior prefrontal sites in representing present option states and of how these state representations are translated into narrower task representations in more posterior sites.

Acknowledgements

This work is supported by the Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/I019847/1, awarded to NY and FW. We thank Kayla Siletti and Michael Clayton for helpful comments on the manuscript.

6 References

Albin, R.L., Young, A.B., Penney, J.B., 1989. The functional anatomy of basal ganglia disorders. Trends in Neurosciences 12(10): 366-375.

Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience 9(1): 357-381.

Allport, A., 1980. Attention and performance, in Claxton, G. (Ed.), Cognitive Psychology: New Directions. Routledge and Kegan Paul, London, pp. 112-153.

Allport, D.A., Styles, E.A., Hsieh, S., 1994. Shifting intentional set: Exploring the dynamic control of tasks, in: Umilta, C., Moscovitch, M. (Eds.), Attention and performance XV. MIT Press, Cambridge, MA, pp. 421-452.

Apicella, P., 2007. Leading tonically active neurons of the striatum from reward detection to context recognition. Trends in Neurosciences 30(6): 299–306.

Aron, A.R., Robbins, T.W., Poldrack R.A.,2014. Inhibition and the right inferior frontal cortex: one decade on. Trends in Cognitive Sciences 18(4): 177–185.

Aron, A.R., L.W., Sahakian, B.J., Monsell, S., Barker, R.A, Robbins, T.W., 2003. Task-Set Switching Deficits in Early-Stage Huntington's Disease: Implications for Basal Ganglia Function. Journal of Cognitive Neuroscience 15(5): 629–642.

Badre, D., D'Esposito, M., 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? Nature Reviews Neuroscience 10(9): 659–669.

Badre, D., Frank, M.J., 2012. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 2: evidence from fMRI. Cerebral Cortex 22(3): 527–536.

Band, G.P.H., Steenbergen, H., Ridderinkhof, K.R., Falkenstein, M., Hommel, B., 2009. Action-effect negativity: irrelevant action effects are monitored like relevant feedback. Biological Psychology 82(3): 211–218.

Barto, A.G., Simsek, O., 2005. Intrinsic motivation for reinforcement learning systems.Proceedings of the Thirteenth Yale Workshop on Adaptive and Learning Systems. New Haven, CT, USA.

Bédard, P., Larochelle, L., Parent, A., Poirier, L.J., 1969. The nigrostriatal pathway: A correlative study based on neuroanatomical and neurochemical criteria in the cat and the monkey. Experimental Neurology 25(3): 365–377.

Bednark, J.G., Reynolds, J.N.J., Stafford, T., Redgrave, P., Franz, E.A., 2013. Creating a movement heuristic for voluntary action: electrophysiological correlates of movement-outcome learning. Cortex 49(3): 771–780.

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. Nature Neuroscience 10(9): 1214–1221.

Bernheimer, H., Birkmayer, W., Hornykiewicz, O., Jellinger, K., Seitelberger, F., 1973. Brain dopamine and the syndromes of Parkinson and Huntington. Clinical, morphological and neurochemical correlations. Journal of the Neurological Sciences 20(4): 415–455.

Berns, G.S., Sejnowski, T.J., 1998. A computational model of how the basal ganglia produce sequences. Journal of Cognitive Neuroscience 10(1): 108–121.

Bischoff-Grethe, A., Crowley, M.G., Arbib, M.A., 2002. Movement inhibition and next sensory state prediction in the basal ganglia, in: Graybiel, A.M., DeLong, M.R., Kitai, S.T. (Eds.), The Basal Ganglia VI. Kluwer Academic/Plenum Publishers, New York. pp. 267–278.

Bjoerklund, A., Dunnett, S.B., 2007. Dopamine neuron systems in the brain: an update. Trends in Cognitive Sciences 30(5): 194–201.

Blakemore, S.-J., Sirigu, A., 2003. Action prediction in the cerebellum and in the parietal lobe. Experimental Brain Research. 153(2): 239–245.

Bogacz, R., Gurney, K., 2007. The basal ganglia and cortex implement optimal decision making between alternative actions. Neural Computation 19(2): 442–477.

Bolam, J.P., Brown, M.T.C., Moss, J., Magill, P.J., 2009. Basal Ganglia : Internal Organization, in: Squire, L.R. (Ed), Encyclopedia of Neuroscience 2. Elsevier/Academic Press. Pp. 97–104.

Botvinick, M., Plaut, D.C., 2004. Doing without schema hierarchies: a recurrent connectionist approach to normal and impaired routine sequential action. Psychological Review 111(2): 395–429.

Botvinick, M.M., 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cognitive, Affective, & Behavioral Neuroscience 7(4): 356–366.

Botvinick, M.M., 2008. Hierarchical models of behavior and prefrontal function. Trends in Cognitive Sciences 12(5): 201–208.

Botvinick, M.M., Niv, Y., Barto, A.C., 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. Cognition 113(3): 262–280.

Cameron, I.G.M., Watanabe, M., Pari, G., Munoz, D.P., 2010. Executive impairment in Parkinson's disease: response automaticity and task switching. Neuropsychologia 48(7): 1948–1957.

Carr, C. (2000). Locating an error correction signal for adult birdsong. Nature Neuroscience 3(5): 419–421.

Cohen, J.D., O'Reilly, R.C., 1996. A preliminary theory of the interactions between pre-frontal cortex and hippocampus that contribute to planning and prospective memory, in: Prospective memory: Theory and applications Brandimonte, M., Einstein, G., McDaniel, M. (Eds.), Erlbaum, Hillsdale, New Jersey. pp. 267–295.

Collins, A.G.E, Cavanagh, J.F., Frank, M.J., 2014. Human EEG uncovers latent generalizable rule structure during learning. The Journal of Neuroscience 34(13): 4677–4685.

Collins, A.G.E., Frank, M.J., 2013. Cognitive control over learning: Creating, clustering, and generalizing task-set structure. Psychological Review 120(1): 190–229.

Collins, A.G.E., Koechlin, E., 2012. Reasoning, Learning, and Creativity: Frontal Lobe Function and Human Decision-Making. PLoS Biol 10(3): e1001293.

Cools, R., 2006. Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. Neuroscience and Biobehavioral Reviews 30(1): 1–23.

Cools, R., Barker, R.A., Sahakian, B.J., Robbins, T.W., 2001. Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. Cerebral Cortex 11(12): 1136–1143.

Cools, R., Barker, R.A., Sahakian, B.J., Robbins, T.W., 2003. L-dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. Neuropsychologia 41(11): 1431–1441.

Cooper, R., Shallice, T., 2000. Contention scheduling and the control of routine activities. Cognitive Neuropsychology 17(4): 297–338.

Crawford, T., Goodrich, S., Henderson, L., Kennard, C., 1989. Predictive responses in Parkinson's disease: manual keypresses and saccadic eye movements to regular stimulus events. Journal of Neurology, Neurosurgery, and Psychiatry 52(9): 1033–1042.

Crittenden, J.R., Graybiel, A.M., 2011. Basal ganglia disorders associated with imbalances in the striatal striosome and matrix compartments. Frontiers in Neuroanatomy 5.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. Neuron 69(6): 1204–1215.

den Ouden, H.E.M., Friston, K.J., Daw, N.D., McIntosh, A.R., Stephan, K.E., 2009. A dual role for prediction error in associative learning. Cerebral Cortex. 19(5): 1175–1185.

Derrfuss, J., Brass, M., Neumann, J., von Cramon, D.Y., 2005. Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. Human Brain Mapping 25(1): 22–34.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annual Review of Neuroscience 18(1): 193–222.

Desrochers, T.M., Badre, D., 2012. Finding parallels in fronto-striatal organization. Trends in Cognitive Sciences 16(8): 407–8.

Di Martino, A., Scheres, A., Margulies, D.S., Kelly, A.M.C., Uddin, L.Q., Shehzad, Z., Biswal, B., Walters, J.R., Castellanos, F.X., Milham, M.P., 2008. Functional connectivity of human striatum: a resting state FMRI study. Cerebral Cortex 18(12): 2735–2747.

Diuk, C., Tsai, K., Wallis, J., Botvinick, M.M., Niv, Y., 2013. Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. The Journal of Neuroscience 33(13): 5797–5805.

Donoso, M., Collins, E.G.E., Koechlin, E., 2014. Foundations of human reasoning in the prefrontal cortex. Science 344(6191): 1481-1485

Dosenbach, N.U.F., Vischer., K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen. S.E., 2006. A core system for the implementation of task sets. Neuron 50(5): 799–812.

Draganski, B., Kherif, F., Klöppel, S., Cook, P.A., Alexander, D.C., Parker, G.J.M., Deichmann, R., Ashburner, J., Frackowiak R.S.J., 2008. Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. The Journal of Neuroscience 28(28): 7143–7152.

Dreisbach, G., Haider, H., 2008. That's what task sets are for: shielding against irrelevant information. Psychological Research 72(4): 355–361.

Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences 14(4): 172–179.

Eblen, F., Graybiel, A.M., 1995. Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. The Journal of Neuroscience 15(9): 5999–6013.

Elliott, R., Frith, C.D., Dolan, R.J., 1997. Differential neural response to positive and negative feedback in planning and guessing tasks. Neuropsychologia 35(10): 1395–1404.

Elsner, B., Hommel, B., 2001. Effect anticipation and action control. Journal of Experimental Psychology: Human Perception and Performance 27(1): 229-240.

Flaherty, A.W., Graybiel, A.M., 1991. Corticostriatal transformations in the primate somatosensory system. projections from physiologically mapped body-part representations. Journal of Neurophysiology 66(4): 1249–1263.

Flaherty, A.W, Graybiel, A.M., 1993. Two input systems for body representations in the primate striatal matrix: experimental evidence in the squirrel monkey. The Journal of Neuroscience 13(3): 1120–1137.

Flowers, K., 1978. Lack of prediction in the motor behaviour of Parkinsonism. Brain, 101(1): 35–52.

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. Journal of Cognitive Neuroscience 17(1): 51–72.

Frank, M.J., 2006. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Networks, 19(8): 1120-1136.

Frank, M.J., Badre, D., 2012. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cerebral Cortex 22(3): 509–526.

Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychological Review 113(2): 300–326.

Frank, M.J., Loughry, B., O'Reilly, R.C., 2001. Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cognitive, Affective, & Behavioral Neuroscience 1(2): 137–160.

Frank, M.J., Samanta, J., Moustafa, A.A., Sherman, S.J., 2007. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science 318(5854): 1309–1312.

Frank, M.J., Scheres, A., Sherman, S.J., 2007. Understanding decision-making deficits in neurological conditions: insights from models of natural action selection. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 362(1485): 1641–1654.

Frank, M.J.,Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306(5703): 1940–1943.

Gaschler, R., Nattkemper, D., 2012. Instructed task demands and utilization of action effect anticipation. Frontiers in Psychology 3(December): 578.

Grahn, J.A., Parkinson, J.A., Owen, A.M., 2008. The cognitive functions of the caudate nucleus. Progress in Neurobiology 86(3): 141–155.

Grahn, J.A., Rowe, J.B., 2013. Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. Cerebral Cortex 23(4): 913–921.

Grange, J.A., Houghton, G., 2014. Task Switching and Cognitive Control : An Introduction, in: Grange J.A., Houghton, D. (Eds.), Task switching and cognitive control, 237. Oxford University Press, Oxford, pp. 1-26.

Graybiel, A.M., 1998. The basal ganglia and chunking of action repertoires. Neurobioly of Learning and Memory 70(1- 2): 119–136.

Graybiel, A.M., 2005. The basal ganglia: learning new tricks and loving it. Current Opinion in Neurobiology 15(6): 638–644.

Graybiel, A.M., 2008. Habits, rituals, and the evaluative brain. Annual Review of Neuroscience 31: 359–87.

Greenwald, A.G., 1970. Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. Psychological Review 77(2): 73-99.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., Passingham, R.E., 2003. Objects automatically potentiate action: an fMRI study of implicit processing. European Journal of Neuroscience 17(12): 2735–2740.

Gurney K.N, Humphries M.D., Redgrave P.,2015. A New Framework for Cortico-Striatal Plasticity: Behavioural Theory Meets In Vitro Data at the Reinforcement-Action Interface. PLoS Biol 13(1): e1002034. doi:10.1371/journal.pbio.1002034

Gurney, K.N., Prescott, T.J., Redgrave, P., 2001. A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biological Cybernetics, 84(6): 401-410.

Haber, S.N., 2003. The primate basal ganglia: parallel and integrative networks. Journal of Chemical Neuroanatomy, 26(4): 317-330.

Haruno, M., Kawato, M., 2006. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. Neural Networks 19(8): 1242–1254.

Hazy, T.E., Frank, M.J., O'Reilly, R.C., 2007. Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. Philosophical Transactions of the Royal Society B: Biological Sciences 362(1485): 1601–1613.

Herwig, A., Prinz, W., Waszak, F., 2007. Two modes of sensorimotor integration in intention-based and stimulus-based actions. Quarterly Journal of Experimental Psychology 60(11): 1540–1554.

Herwig, A., Waszak, F., 2009. Intention and attention in ideomotor learning. Quarterly Journal of Experimental Psychology 62(2): 219–227.

Hikosaka, O., Bromberg-Martin, E., Hong, S., Matsumoto, M., 2008. New insights on the subcortical representation of reward. Current Opinion in Neurobiology 18(2): 203–208.

Holl, A.K., Wilkinson, L., Tabrizi, S.J., Painold, A., Jahanshahi, M., 2012. Probabilistic classification learning with corrective feedback is selectively impaired in early Huntington's disease—evidence for the role of the striatum in learning with feedback. Neuropsychologia 50(9): 2176–2186.

Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychological Review 109(4): 679–709.

Holroyd, C.B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R.B., Coles, M.G.H., Cohen, J.D., 2004. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. Nature Neuroscience 7(5): 497–498.

Holroyd, C.B., Yeung, N. 2012. Motivation of extended behaviors by anterior cingulate cortex. Trends in Cognitive Sciences 16(2): 122–128.

Hommel, B., 2009. Action control according to TEC (theory of event coding). Psychological Research 73(4): 512–526.

Horvitz, J.C., 2000. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience 96(4): 651–656.

Houk, J.C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P.J.R., Roy, S.A., Simo, L.S., 2007. Action selection and refinement in subcortical loops through basal ganglia and cerebellum. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 362(1485): 1573–1583.

Houk, J.C., Wise, S.P., 1995. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cerebral Cortex 5(2): 95–110.

Ito, J., Kitagawa, J., 2006. Performance monitoring and error processing during a lexical decision task in patients with Parkinson's disease. Journal of Geriatric Psychiatry and Neurology 19(1): 46–54.

Jahfari, S., Verbruggen, F., Frank, M.J., Waldorp, L.J., Colzato, L., Ridderinkhof, K.R., Forstmann B.U., 2012. How preparation changes the need for top-down control of the basal ganglia when inhibiting premature actions. The Journal of Neuroscience 32(32): 10870–10878.

James, W., 1890. The principles of psychology. MacMillan, London, UK

Janczyk, M., Alexander H., Pfister, R., 2012. Instant attraction: immediate action-effect bindings occur for both, stimulus-and goal-driven actions. Frontiers in Psychology 3(October): 446.

Jersild, A.T., 1927. Mental set and shift. Archives of Psychology, 89.

Jin, X., Tecuapetla, F., Costa, R.M., 2014. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. Nature Neuroscience 17: 423–430

Joel, D., Niv, Y., Ruppin, E., 2002. Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Networks 15(4-6): 535–547.

Joel, D., Weiner, I., 2000. 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(3): 451-474.

Jueptner, M., Weiller, C., 1998. A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. Brain 121(8): 1437–1449.

Kemp, J.M., Powell, T.P.S., 1970. The cortico-striate projection in the monkey. Brain 93(3): 525–546.

Keramati, M., Gutkin, B., 2013. Imbalanced Decision Hierarchy in Addicts Emerging from Drug-Hijacked Dopamine Spiraling Circuit. PLoS ONE 8(4): e61489.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A.M., Koch I., 2010. Control and interference in task switching—a review. Psychological Bulletin 136(5): 849-874.

Kim, C., Cilles, S.E., Johnson, N.F., Gold, B.T., 2012. Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. Human Brain Mapping 33(1): 130–142.

Klein, T.A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D.Y., Ullsperger, M., 2007. Genetically determined differences in learning from errors. Science 318(5856): 1642–1645.

Kleinsorge, T., Rinkenauer G., 2012. Effects of monetary incentives on task switching. Experimental Psychology 59(4): 216-226.

Knowlton, B.J., Squire, L.R., Gluck, M.A., 1994. Probabilistic classification learning in amnesia. Learning & Memory 1(2): 106–120.

Koechlin, E., Hyafil, A., 2007. Anterior prefrontal function and the limits of human decision-making. Science 318(5850): 594–598.

Koechlin, E., Ody, C., Kouneiher, F., 2003. The architecture of cognitive control in the human prefrontal cortex. Science 302(5648): 1181–1185.

Koechlin, E., Summerfield, C., 2007. An information theoretical approach to prefrontal executive function. Trends in Cognitive Sciences 11(6): 229–235.

Krieghoff , V., Waszak, F., Prinz, W., Brass, M., 2011. Neural and behavioral correlates of intentional actions. Neuropsychologia 49(5): 767–776.

Krigolson, O.E., Holroyd, C.B., 2006. Evidence for hierarchical error processing in the human brain. Neuroscience 137(1): 13–17.

Kringelbach, M.L., O'Doherty, J.P., Rolls, E.T., Andrews, C., 2003. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. Cerebral Cortex 13(10): 1064–1071.

Kunde, W., 2001. Response-Effect Compatibility in Manual Choice Reaction Tasks. Journal of Experimental Psychology 27(2): 387-394

Kühn, S., Reurinck, R., Fias, W., Waszak, F., 2010. The Internal Anticipation of Sensory Action Effects: When Action Induces FFA and PPA Activity. Frontiers in Human Neuroscience 4(June): 54.

Kowall, N.W., Ferrante, R.J., Martin, J.B., 1987. Patterns of cell loss in Huntington's disease. Trends in Neurosciences 10(1): 24-29.

Lashley, K.S., 1951. The problem of serial order in behavior. in: Jeffress, L.A. (Ed.). Cerebral mechanisms in behaviour: The Hinxon Symposium. Wiley & Sons, New York. pp. 112–135.

Lehericy, S., Benali, H., Van de Moortele, P.F., Pelegrini-Issac, M., Waechter, T., Ugurbil, K., Doyon, J., 2005. Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. PNAS, 102(35): 12566–12571.

Lukas, S., Philipp, A.M., Koch, I., 2012. The influence of action effects in task-switching. Frontiers in Psychology 3(January): 595.

MacLeod, C.M., 1991. Half a century of research on the Stroop effect: an integrative review. Psychological Bulletin 109(2): 163.

Marien, H., Aarts, H., Custers, R., 2012. Being flexible or rigid in goal-directed behavior: When positive affect implicitly motivates the pursuit of goals or means. Journal of Experimental Social Psychology 48(1): 277–283.

Marien, H., Aarts, H., Custers, R., 2013. Adaptive control of human action: the role of outcome representations and reward signals. Frontiers in Psychology 4(September): 602.

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

Matsumoto, N., Hanakawa, T., Maki, S., Graybiel, A.M., Kimura, M., 2009. Nigrostriatal Dopamine System in Learning to Perform Sequential Motor Tasks in a Predictive Manner. Journal of Neurophysiology 82(2): 978–998.

Meiran, N., 1996. Reconfiguration of processing mode prior to task performance. Journal of Experimental Psychology: Learning, Memory, and Cognition 22(6): 1423-1442.

Miall, R.C., 2003. Connecting mirror neurons and forward models. Neuroreport 14(17): 2135–2137.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annual Review of Neuroscience 24: 167–202.

Miller, G.A., Galanter, E., Pribram, K.H., 1960. Plans and the structure of behavior. Holt, Rinehart and Winston. Inc., New York.

Monchi, O., Petrides, M., Petre, V., Worsley, K., Dagher, A., 2001. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. The Journal of Neuroscience 21(19): 7733–7741.

Monsell S., 1996. Control of mental processes, in: Bruce, V. (Ed.), Unsolved mysteries of the mind. Erlbaum, Hove, E. Sussex, pp. 93-148.

Monsell, S., Driver, J., 2000. Control of Cognitive Processes. MIT Press, Cambridge, MA.

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

Morris, G., Schmidt, R., Bergman, H., 2010. Striatal action-learning based on dopamine concentration. Experimental Brain Research 200(3-4): 307–317.

Nambu, A., 2004. A new dynamic model of the cortico-basal ganglia loop. Progress in Brain Research 143: 461–466.

Nambu, A., Hironobu, T., Takada, M., 2002. Functional significance of the cortico– subthalamo–pallidal ‘hyperdirect’pathway. Neuroscience Research 43(2): 111–117.

Nieuwenhuis, S., Heslenfeld, D.J., Alting, N.J., Geusau, V., Mars, R.B., Holroyd, C.B., Yeung N., 2005. Activity in human reward-sensitive brain areas is strongly context dependent. NeuroImage 25(4): 1302–1309.

Nieuwenhuis, S., Monsell, S., 2002. Residual costs in task switching: Testing the failure-to-engage hypothesis. Psychonomic Bulletin & Review 9(1): 86–92.

Nieuwenhuis, S., Slagter, H.A., Geusau, V., Alting, N.J., Heslenfeld, D.J., Holroyd, C.B., 2005. Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. European Journal of Neuroscience 21(11): 3161–3168.

Nobili, F., Campus, C., Arnaldi, D., De Carli, F., Cabassi, G. Brugnolo, A., Dessi, B., Morbelli, S., Sambuceti, G., Abbruzzese, G., Rodriguez G., 2010. Cognitive nigrostriatal relationships in de novo, drug-naïve Parkinson’s disease patients: a [I-123]FP-CIT SPECT study. Movement Disorders 25(1): 35–43.

Norman D.A., Shallice, T., 1986. Attention to action: Willed and automatic control of behaviour, in: Davidson, R.J., Schwartz, G.E., Shapiro, D. (Eds.), Consciousness and self-regulation, vol 4. Plenum, New York, pp. 1-18

O'Doherty, J. P., 2004. Reward representations and reward-related learning in the human brain: insights from neuroimaging. Current Opinion in Neurobiology 14(6): 769–776.

O'Reilly, R.C., 2006. Biologically based computational models of high-level cognition. Science 314(5796): 91–94.

O'Reilly, R.C., Frank, M.J., 2006. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural computation 18(2): 283–328.

Parent, A., Hazrati, L., 1995. Functional anatomy of the basal ganglia . II . The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. Brain Research Reviews 20: 128–154.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442(7106): 1042–1045.

Prinz, W., 1990. A common coding approach to perception and action, in: Neumann, O., Prinz, W. (Eds.), Relationships between perception and action. Springer-Verlag, Berlin, pp. 167-201.

Redgrave, P., Gurney, K.N., 2012. The short-latency dopamine signal: a role in discovering novel actions? Nature Reviews Neuroscience 7: 967-975.

Redgrave, P., Prescott, T.J., Gurney, K., 1999. The basal ganglia: a vertebrate solution to the selection problem?. Neuroscience, 89(4): 1009-1023.

Reynolds, J.N.J., Wickens, J.R., 2002. Dopamine-dependent plasticity of corticostriatal synapses. Neural Networks 15(4): 507-521.

Reynolds, J.N.J, Hyland, B.I., Wickens, B.R., 2001. A cellular mechanism of reward-related learning. Nature 413(6851): 67–70.

Ribas-Fernandes, J.J.F., Solway, A., Diuk, C., McGuire, J.T., Barto, A.G, Niv, Y., Botvinick, M.M., 2011. A neural signature of hierarchical reinforcement learning. Neuron 71(2): 370–379.

Richter, F., Yeung, N., 2014. Neuroimaging studies of task switching, in: Grange J.A., Houghton, D. (Eds.), Task switching and cognitive control. Oxford University Press, Oxford, pp. 237 – 271.

Riederer, P., Wuketich, S., 1976. Time course of nigrostriatal degeneration in Parkinson's disease. Journal of Neural Transmission 38(3-4): 277-301.

Ridderinkhof, K. R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S., 2004. The role of the medial frontal cortex in cognitive control. Science 306(5695): 443–447.

Robbins, T.W., James, M., Owen, A.M., Lange, K.W., Lees, A.J., Leigh, P.N., Marsden, C.D., Quinn, N.P., Summers, B.A., 1994. Cognitive deficits in progressive supranuclear palsy, Parkinson's disease, and multiple system atrophy in tests sensitive to frontal lobe dysfunction. Journal of Neurology, Neurosurgery & Psychiatry 57(1): 79–88.

Rogers, R.D., Sahakian, B.J., Hodges, J.R., Polkey, C.E., Kennard, C., Robbins T.W., 1998. Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. Brain 121(5): 815–842.

Rogers, R.D., Monsell, S., 1995. Costs of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology : General 124(2): 207-231

Ruge, H., Müller, S.C., Braver, T.S.,2010. Anticipating the consequences of action: An fMRI study of intention-based task preparation. Psychophysiology, 47: 1019-1027

Saint-Cyr, J.A., 2003. Frontal-striatal circuit functions: context, sequence, and consequence. Journal of the International Neuropsychological Society 9(1): 103–127.

Sakai, K., 2008. Task set and prefrontal cortex. Annual Review of Neuroscience 31: 219–245.

Sardo, P., Ravel, S., Legallet, E., Apicella, P., 2000. Influence of the predicted time of stimuli eliciting movements on responses of tonically active neurons in the monkey striatum. European Journal of Neuroscience 12(5): 1801–1816.

Schaal, S., Mohajerian, P., Ijspeert, A., 2007. Dynamics systems vs. optimal control–a unifying view. Progress in brain research 165(1): 425–445.

Schacter, D.L., Addis, D.R., Buckner, R.L., 2007. Remembering the past to imagine the future: the prospective brain. Nature Reviews Neuroscience 8(9): 657–661.

Schiffer, A.-M., Ahlheim, C., Wurm, M.F., Schubotz, R.I., 2012. Surprised at all the entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. PloS one 7(5):e36445.

Schiffer, A.-M., Ahlheim, C., Ulrichs, K., Schubotz R.I., 2013. Neural changes when actions change: Adaptation of strong and weak expectations. Human Brain Mapping 34(7): 1713 - 1727

Schiffer, A.-M., Schubotz, R.I., 2011. Caudate nucleus signals for breaches of expectation in a movement observation paradigm. Frontiers in human neuroscience 5: 38.

Schmahmann, J.D., Pandya, D.N., 2006. Fiber Pathways of The Brain. Oxford University Press, Oxford.

Schmahmann, J.D., Pandya, D.N., 2008. Disconnection syndromes of basal ganglia, thalamus, and cerebrocerebellar systems. Cortex 44(8): 1037–1066.

Schneider, D.W., Logan, G.D., 2006. Hierarchical control of cognitive processes: Switching tasks in sequences. Journal of Experimental Psychology: General 135(4): 623-640

Schultz, W., 2000. Multiple reward signals in the brain. Nature reviews neuroscience 1(3): 199–207.

Schultz, W., Dickinson, A., 2000. Neuronal coding of prediction errors. Annual review of neuroscience 23: 473–500.

Schultz, W., 2007. Behavioral dopamine signals. Trends in neurosciences 30(5): 203–210.

Schultz, W., 2013. Updating dopamine reward signals. Current opinion in neurobiology 23(2): 229–238.

Schultz, W., Dayan, P., Montague, P.R., 1997. A Neural Substrate of Prediction and Reward. Science 275(June 1994): 1593–1599.

Schultz, W., Apicella, P., Scarnati, E., Ljungberg, T., 1992. Neuronal activity in monkey ventral striatum related to the expectation of reward. The Journal of Neuroscience, 12(12): 4595-4610.

Seger, C.A., 2008. How do the basal ganglia contribute to categorization? their roles in generalization, response selection, and learning via feedback. Neuroscience & Biobehavioral Reviews 32(2): 265–278.

Seger, C.,A., Spiering, B.J., Sares, A.G., Quraini, S.I., Alpeter, C., David, J., Thaut, M.H., 2013. Corticostriatal contributions to musical expectancy perception. Journal of cognitive neuroscience 25(7): 1062–1077.

Seitz, A.R., Kim, D., Watanabe, T., 2009. Rewards evoke learning of unconsciously processed visual stimuli in adult humans. Neuron 61(5): 700–707.

Selemon, L.D., Goldman-Rakic, P.S., 1985. Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. The Journal of Neuroscience. 5(3): 776–794.

Shen, Y.J., Chun, M.M., 2011. Increases in rewards promote flexible behavior. Attention, perception & psychophysics 73(3): 938–52.

Shohamy, D., Myers, C.E., Kalanithi, J., Gluck, M.A., 2008. Basal ganglia and dopamine contributions to probabilistic category learning. Neuroscience and biobehavioral reviews 32(2): 219–236.

Shook, S.K., Franz, E.A., Higginson, C.I., Wheelock, V.L., Sigvardt, K.A., 2005. Dopamine dependency of cognitive switching and response repetition effects in Parkinson's patients. Neuropsychologia 43(14): 1990–1999.

Smith, Y., Beyan, M.D., Shink, E., Bolam, J.P., 1998. Microcircuitry of the direct and indirect pathways of the basal ganglia. Neuroscience 86: 353-388.

Smith, Y., Raju, D.V., Pare, J.-F., Sidibe, M., 2004. The thalamostriatal system: a highly specific network of the basal ganglia circuitry. Trends in Neurosciences 27(9): 520–527.

Stalnaker, T.A., Cooch, N.K., McDannald, M.A., Liu, T.-L., Wied, H., Schoenbaum, G., 2014. Orbitofrontal neurons infer the value and identity of predicted outcomes. Nature Communications 5, 3962

Stocco, A., Lebiere, C., Anderson, J.R., 2010. Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. Psychological review 117(2): 541–574.

Suri, R.E., 2002. TD models of reward predictive responses in dopamine neurons. Neural Networks 15(46): 523–533.

Sutton, R.S., Barto, A.G., 1990. Time-derivative models of Pavlovian reinforcement, in: Gabriel, M., Moore, J. (Eds.), Learning and Computational Neuroscience: Foundations of Adaptive Networks, MIT Press, Cambridge, MA, pp. 497–537.

Sutton, R.S., 1988. Learning to predict by the methods of temporal differences. Machine learning 3(1): 9–44.

Sutton, R.S., Barto, A.G., 1981. An adaptive network that constructs and uses an internal model of its world. Cognition and Brain Theory 4(3): 217–246.

Swanson DePasque, S., Tricomi, E., 2014. Goals and task difficulty expectations modulate striatal responses to feedback. Cognitive, Affective, & Behavioral Neuroscience 14: 610-620.

Thorndike, E.L., 1927. The Law of Effect. The American Journal of Psychology 39(1): 212–222.

Torrecillos, F., Albouy, P., Brochier, T., Malfait, N., 2014. Does the processing of sensory and reward-prediction errors involve common neural resources? Evidence from a frontocentral negative potential modulated by movement execution errors. The Journal of Neuroscience 34(14): 4845–4856.

Tricomi, E., Fiez, J.A., 2008. Feedback signals in the caudate reflect goal achievement on a declarative memory task. NeuroImage 41(3): 1154–1167.

Tunik, E., Houk, J.C., Grafton, S.T., 2009. Basal ganglia contribution to the initiation of corrective submovements. NeuroImage 47(4): 1757–1766.

Ullsperger, M., von Cramon, D.Y., 2003. Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. The Journal of Neuroscience 23(10): 4308-4314.

Ullsperger, M., von Cramon, D.Y., 2006. The role of intact frontostriatal circuits in error processing. Journal of Cognitive Neuroscience 18(4): 651-664.

van Schouwenburg, M.R., O'Shea, J., Mars, R.B., Rushworth, M.F.S., Cools, R., 2012. Controlling human striatal cognitive function via the frontal cortex. The Journal of Neuroscience 32(16): 5631-5637.

Vandierendonck, A., Liefooghe, B., Verbruggen F., 2010. Task switching: interplay of reconfiguration and interference control. Psychological bulletin 136(4): 601-626.

Walsh, M.M., Anderson, J.R., 2012. Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. Neuroscience and biobehavioral reviews 36(8): 1870–1884.

Waszak, F., Wascher, E., Keller, P., Koch, I., Aschersleben, G., Rosenbaum, D.A., Prinz, W., 2005. Intention-based and stimulus-based mechanisms in action selection. Experimental brain research 162(3): 346–356.

Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., Niv, Y., 2014. Orbitofrontal cortex as a cognitive map of task space. Neuron 81(2): 267-279.

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

Wolpert, D.M., Kawato, M.,1998. Multiple paired forward and inverse models for motor control. Neural Networks 11(7-8): 1317–1329.

Wolpert, D.M., Miall, R.C., 1996. Forward Models for Physiological Motor Control. Neural Networks 9(8): 1265–1279.

Yeung, N., Sanfey, A.G., 2004. Independent coding of reward magnitude and valence in the human brain. The Journal of Neuroscience 24(28): 6258–6264.

Yeung, N., 2013. Conflict monitoring and cognitive control, in: Oxford handbook of cognitive neuroscience, Volume 2, Ochsner, K., Kosslyn, S. (Eds.), Oxford University Press, Oxford, pp. 275–299

Yousry, T.A., Schmid, U.D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., Winkler, P., 1997. Localization of the motor hand area to a knob on the precentral gyrus. a new landmark. Brain 120(1): 141–157.

Zalla, T., Sirigu, A., Pillon, B., Dubois, B., Grafman, J., Agid, Y., 1998. Deficit in Evaluating Pre-Determinated Sequences of Script Events in Patients with Parkinson's Disease. Cortex 34(4): 621–627.

Ziessler, M., Nattkemper, D., Vogt, S., 2012. The activation of effect codes in response preparation: new evidence from an indirect priming paradigm. Frontiers in Psychology 3: 585.