



Published in final edited form as:

Neural Netw. 2009 ; 22(5-6): 757–765. doi:10.1016/j.neunet.2009.06.008.

Computational Perspectives on Forebrain Microcircuits Implicated in Reinforcement Learning, Action Selection, and Cognitive Control

Daniel Bullock *

Boston University, Department of Cognitive and Neural Systems, 677 Beacon Street, Boston, MA 02215

Can Ozan Tan, and

Harvard Medical School, Boston, MA Spaulding Rehabilitation Hospital, 125 Nashua Street, Boston, MA 02114

Yohan J. John

Boston University, Department of Cognitive and Neural Systems, 677 Beacon Street, Boston, MA 02215

Abstract

Abundant new information about signaling pathways in forebrain microcircuits presents many challenges, and opportunities for discovery, to computational neuroscientists who strive to bridge from microcircuits to flexible cognition and action. Accurate treatment of microcircuit pathways is especially critical for creating models that correctly predict the outcomes of candidate neurological therapies. Recent models are trying to specify how cortical circuits that enable planning and voluntary actions interact with adaptive sub-cortical microcircuits in the basal ganglia. The basal ganglia are strongly implicated in reinforcement learning, and in all behavior and cognition over which the frontal lobes exert flexible control. The persisting role of the basal ganglia shows that ancient vertebrate designs for motivated action-selection proved adaptable enough to support many “modern” behavioral innovations, including fluent generation of language and speech. This paper summarizes how recent models have incorporated realistic representations of microcircuit features, and have begun to trace their computational implications. Also summarized are recent empirical discoveries that provide guidance regarding how to formulate the rules for synaptic modification that govern learning in cortico-striatal pathways. Such efforts are contributing to an emerging synthesis based on an interlocking set of computational hypotheses regarding cortical interactions with basal ganglia and thalamic nuclei. These hypotheses specify how specialized microcircuits solve learning and control problems inherent to the brain's parallel design.

Keywords

Basal ganglia; acetylcholine; dopamine; striatum; decision making

© 2009 Elsevier Ltd. All rights reserved.

*Corresponding author *Email addresses:* danb@cns.bu.edu (Daniel Bullock), cotan@partners.org (Can Ozan Tan), yohan@cns.bu.edu (Yohan J. John).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction: Forebrain Circuits for Motivated Action Selection

A recurring task for an animal is to select, among probably-achievable action plans, those action plans that are more likely to promote its well-being. Thus, the planner needs access to frequently updated estimates not only of act-outcome probabilities and outcome values, but also of the achievability of the actions, given both the actor's current state and the context of action. Because plan evaluation and selection may occur well before the best time to execute a plan, and because plans take time to execute, new options may be noticed after plan selection, but before plan completion. Thus, it will often pay to interrupt execution of a plan, perform another, and then resume or abandon the original plan. Coping well with such complications requires sophistication in the forebrain circuitry that enables intelligent planning in mammals.

Although the forebrain encompasses the cerebral cortex and BG (basal ganglia), most treatments of intelligent planning and motivated action–selection validly focus on the frontal cortex and BG, which are so linked that the BG exert a much stronger and more direct influence on frontal than on posterior cortex. Indeed, a general rule is that for every planning and action deficit that results from lesioning a discrete region within frontal cortex, a highly similar deficit can be produced by lesioning whatever discrete part of the BG circuit projects its output toward the given region within frontal cortex. This makes sense on the hypothesis, now broadly supported (Bullock, 2004b), that frontal cortex can represent many potential types of cognitive and other actions, whereas the BG are responsible for selecting which such actions to execute (Redgrave et al., 1999a).

Given that selection among action plans is strongly influenced by reinforcement-guided learning, it is not surprising that the BG play multiple roles in action choice and learning. At the coarse grain of a tri-partite division of the conventional BG's major input nucleus, the striatum, it is possible to distinguish: the ventral striatum as important for processing context-conditioned reward expectations, medial-dorsal striatum for act-outcome expectations, and lateral-dorsal striatum for habitual condition action relations. Two of these divisions correspond to an actor-critic architecture (Houk et al., 1995), with learning in both actor and critic compartments governed in part by reward-prediction errors (RPEs), as expected by temporal difference (TD) models of reinforcement learning (e.g., Sutton and Barto, 1981). However, many microcircuit specializations in the BG go beyond or deviate from what was predicted by actor-critic architectures, or by associated TD models, and their existence raises the problem of identifying their computational implications. Moreover, Swanson (2000) has argued for extending the concept of the BG to encompass further regions, such as some nuclei of the extended amygdala. Any such extension would entail further differentiation of architectures and models. In this review, we detail BG microcircuit features, particularly in striatum, which suggest revisions to actor-critic systems and TD-based learning rules; summarize some of their computational implications; and note in-depth treatments published elsewhere. A briefer treatment of these issues appeared in Bullock and Tan (2009).

2. Microcircuit Specializations of the Basal Ganglia, with Computational Implications

2.1. Dopamine (DA) cell firing provides an asymmetrical representation of positive and negative reward prediction errors (RPEs)

Dopamine cells are well established as carriers of conditioned (i.e., stimulus- and learning-history-dependent) RPE signals. That they discharge spontaneously at a constant rate means that positive rate deviations can signal positive RPEs and negative deviations (dips below baseline rate) negative RPEs. However, the low value of the baseline firing rate implies a truncation of negative RPE signals, whereas some DA burst signals to positive RPEs remain

proportionate. For a general class of dual-path models capable of learning to compute RPEs in a network including DA cells, Tan et al. (2008) derived a key computational implication of this asymmetry, which accords with DA signal measurements reported by Tobler et al. (2005). Tan et al. (2008) showed that this asymmetry reflects incremental learning that, after learning, the magnitude of residual phasic DA burst signals generated in response to delivery of a CS-predicted primary reward will scale with the learned conditional probability of reward omission (i.e., $1 - p(R^*|CS)$) given the predictive cue (CS). Note that in their derivation, as in the data (Tobler et al., 2005), the residual DA burst signal is independent of the absolute reward magnitude, R^* . Yet the dopamine burst signal generated earlier, at CS-onset, reflects both R^* and $p(R^*|CS)$. Although the concept of RPEs is helpful for interpreting DA burst signals, any interpretation that equates the two is incorrect.

2.2. Conditional components of DA signaling go well beyond RPEs, to include most (all?) of what is expected from behavioral reinforcement research

Behavioral research on what constitutes a reinforcer has shown that, in addition to prototypical events, such as reward delivery, several other types of events behave as reinforcers, i.e., such events' contingent delivery raises the future probability of cue-conditioned instrumental behavior. Notable cases are: contingent cessation of an aversive input, contingent delivery of non-aversive stimuli that are novel but not linked to tangible rewards, and contingent access to the opportunity to engage in a more preferred behavior. Of these, the first two have been shown to have the effect on DA release that would be expected if such release also mediates these components of internal reinforcement signaling (see also Ungless et al., 2004). Some initial computational implications were disclosed using the new striatal microcircuit model of Tan and Bullock (2008b). The novelty-related DA cell responses present a further challenge to the classical view of DA signals as pure RPE signals. Such novel events are neither conventional rewards nor conventional learned predictors of such. Recently, two studies attempted to address the novelty-related increase in DA release within a formal TD framework. Kakade and Dayan (2002) assumed that novelty, in itself, provides "reward bonuses" that alter the sum of experienced reward, thereby interfering with RPE estimations with respect to normal external rewards. In an attempt to avoid treating novelty as inherently rewarding, Laurent (2008) presented a TD-based simulation that showed how positive RPEs to novel events (state entries) might arise from prior reinforcement learning. Although it seems likely that some of the ability of novel stimuli to generate DA responses results from their similarity to other reward-predicting cues, the treatment in Laurent (2008) does not appear to be compatible with the key observation that the novelty-related DA responses habituate as a cue's novelty wanes (e.g., Redgrave et al., 1999b). Note that novelty wanes as the cognitive system learns correct predictions in general - not just correct predictions about rewards. Again, DA burst signals cannot be accurately modeled from a perspective that focuses exclusively on RPEs.

2.3. Sustained dopamine cell signal components reflect uncertainty in reward prediction

Contrary to expectations of TD models, there is evidence (Fiorillo et al., 2003, 2005) that DA cells exhibit an uncertainty response that is a non-monotonic (inverted-U) function of reward probability conditional on the cue: when there is a probabilistic relationship between a predictive cue and a rewarding outcome that may (or may not) occur at a fixed time after cue onset, then there is a gradual buildup of DA cell firing rate between the cue onset and the expected time of the uncertain reward. In an attempt to derive an uncertainty response within the TD framework Niv et al. (2005) proposed that this response emerges as an artifact of averaging *gradually* back-propagating RPEs across successive learning trials. To the contrary, there is no evidence that RPE signals "propagate *gradually* backwards" in the time between cue and reward, and the data (Fiorillo et al., 2003, 2005) show that uncertainty responses are robust on single trials. Uncertainty response appear to be inexplicable using a standard TD model.

More recently, Tan and Bullock (2008b,c) proposed that this signal component may be computed by a surprisingly common but rarely simulated property of neurons: co-release of more than one chemical signaler from the same axon terminal. They showed that the well-established, but computationally mysterious, co-release of GABA and the neuropeptide SP (substance P) from striato-nigral terminals can explain robust single trial computation of uncertainty responses, which are a non-monotonic function of the conditional probability of a reward (R^*) given a cue (CS), i.e., $p(R^*|CS)$. Under broad conditions, such co-release will produce a signal proportional to $p(1-p)$, which shows a peak when uncertainty is maximal, i.e., when $p(R^*|CS) = .5$. Although the discoverers of this DA signal component aptly proposed that it may be important to explain habitual gambling, Tan and Bullock (2008b) noted that the broadcast of the DA signal to many brain sites beyond the dorsolateral striatum implies that it can function much more broadly, and adaptively, to optimize computations in both learning and performance. Notably, it can promote search for more-predictive representations, and rapid switching away from no-longer-rewarding alternatives. This functional interpretation of the role of sustained elevation of DA level and its origins (co-release of SP from striato-nigral terminals) deviates significantly from simple RPE and TD frameworks, but provides an explanation for the behavioral effects of SP: injection of SP into the VTA enhances responding for conditioned rewards in general, but also disrupts reward discrimination processes and thereby results in (some degree of) response generalization (Placenza et al., 2004; Kelley et al., 1989).

Such specialized task-dependent firing patterns of the DA cells highlight shortcomings of the computational models of DA responses that are based on the formal RPE-TD framework. The genesis and effects of at least three distinct DA cell firing patterns (reward-related bursts, novelty responses, and sustained uncertainty responses), which exhibit distinct task-dependencies and operate on different time-scales, suggest that DAergic projections to striatal target structures (in particular, to dorsal striatum) engender computations that go beyond those envisioned in the RPE hypothesis of DA and the “actor-critic” concept of BG architecture. In addition, oft-neglected interactions between neurotransmitters, coupled with specialized microcircuits of the BG, indicate a need to revise common reinforcement rules, and imply a far-reaching role for the BG, not only in reinforcement learning but also in evaluation, selection, and execution of actions whose outcomes are contingent on diverse types of factors. We discuss neurotransmitter interactions next, and BG micro- and macrocircuits in Section 2.7

2.4. Striatal learning, adaptive timing and action-gating are governed by a dopamine-acetylcholine (DA-ACh) cascade

Early studies of Parkinson's disease emphasized the striatal balance between DA and ACh in the performance-control functions of the striatum. Because the only sources for striatal ACh are the giant ACh neurons of the striatum itself, the striatal ACh signal source is anatomically distinct from the ACh cells whose projections strongly affect attention/arousal as well as neocortical (Kilgard and Merzenich, 1998) and hippocampal (Hasselmo, 2006) learning. Striatal ACh neurons are often called TANs (for “tonically active neurons”), and their functional signaling obeys different principles than non-striatal ACh neurons. Like DA cells, they show learning-dependent changes. Recently Tan and Bullock (2008a) showed, in a biophysically realistic simulation of TANs, that many of these learning-dependent changes are attributable to learning-dependent changes in the behavior of DA cells whose axons synapse on the TANs. Because both DA and ACh modulate learning of cortico-striatal synapses (Centonze et al., 1999; Pawlak and Kerr, 2008; Wang et al., 2006), it is now clear that both striatal learning and striatal performance functions are strongly dependent on a DA-ACh cascade. One immediate implication is that common reinforcement learning rules need updating to reflect an additional ACh dependency.

Many common reinforcement learning rules have been solely based on DAergic reward prediction errors (RPEs). These rules are often used in combination with the concept of direct and indirect pathways in the BG, (e.g., Albin et al., 1989). According to this scheme, cortical signals are distributed to two classes of striatal output neurons (medium spiny projection neurons; MSPN). MSPNs that contain neuropeptide substance-P (SP) and express mainly D1-type DA receptors (D1-SP-MSPNs hereafter) make direct contact with the BG output nuclei, forming the *direct pathway*. MSPNs that contain enkephalin and express mainly D2-type DA receptors (D2-ENK-MSPNs hereafter) contact BG output nuclei indirectly via relays in the globus pallidus and STN (subthalamic nucleus), forming the *indirect pathway*. The direct pathway is assumed to promote or permit behaviors (the “GO” pathway), whereas the indirect pathway is assumed to suppress or inhibit behaviors (the “NO-GO” or “STOP” pathway). Reinforcement learning rules for acquiring behaviors in this simplified system posit a D1 receptor-mediated long-term potentiation (LTP) of cortico-striatal synapses onto the direct pathway MSPNs, and D2 receptor-mediated long-term depression (LTD) of cortico-striatal synapses onto the indirect pathway MSPNs. Therefore, phasic DA signals (presumed to reflect RPEs, but see above) are assumed to drive learning in opposite directions in these two pathways (e.g. Frank, 2005; Brown et al., 2004). This presumption is based on the earlier observation that LTD and LTP occur at the synapses between cortical pyramidal cells and striatal MSPNs (Calabresi et al., 1992a,b), and that dopaminergic D2 (and to some extent, D1) receptors are crucial for LTD induction (Calabresi et al., 1992a; Kerr and Wickens, 2001), whereas induction of LTP depends critically on the D1 dopamine receptors (Kerr and Wickens, 2001; Schotanus and Chergui, 2008). However, as briefly mentioned above, a growing body of evidence contradicts this simple (yet convenient) learning rule: (1) ACh strongly modulates striatal plasticity via muscarinic receptors, and (2) both LTD and LTP can occur at synapses onto both D1-receptor bearing and D2-receptor bearing MSPN classes.

Both *in vitro* pharmacological and *in vivo* gene-knockout studies have shown that a pause in the striatal ACh signal is required for LTD at corticostriatal synapses onto both classes of MSPNs, whereas baseline or elevated cholinergic transmission is necessary for LTP (Centonze et al., 1999; Bonsi et al., 2008; Wang et al., 2006). More specifically, Bonsi et al. (2008) showed that either pharmacologic blockade or genetic deletion of muscarinic M₂/M₄ receptors, which serve as autoreceptors on TANs that limit striatal ACh level, impairs LTD but not LTP, and this impairment is alleviated by either depleting striatal ACh or blocking postsynaptic M₁ receptors, which are located on MSPNs. Although LTP induction was unaffected by blockade/deletion of presynaptic M₂/M₄ autoreceptors, activation of postsynaptic M₁ receptors is necessary for LTP induction. In fact, in the presence of an M₁ receptor antagonist, cortical high-frequency stimulation failed to induce LTP on the recipient MSPNs, even when the dopamine D2 receptors were blocked concomitantly. This observation suggests that the lack of LTP induction was not due to interference by the dopamine D2 receptors, and confirms a role for postsynaptic muscarinic M₁ receptors in LTP induction. In summary, it appears that LTD requires intact presynaptic M₂/M₄ autoreceptors, or a reduction of the striatal ACh signal to below-baseline levels, whereas LTP induction requires stimulation of postsynaptic M₁ by baseline or elevated ACh transmission.

Spike timing dependent plasticity (STDP) adds another piece to the puzzle of cortico-striatal learning. Rules of cortico-striatal STDP for synapses onto MSPNs are of reversed direction (Fino et al., 2005) compared to those abstracted from observations on other brain structures (e.g., Sjöström and Nelson, 2002). That is, LTP occurs when a postsynaptic MSPN is activated *before* cortical high frequency stimulation (“post-pre” LTP), whereas LTD is observed when a postsynaptic MSPN is activated *after* cortical stimulation (“pre-post” LTD). In addition, while this study by Fino et al. (2005) did not identify MSPN classes, they reported that bidirectional plasticity (i.e., both LTP and LTD) occurs at the same cortico-striatal synapses, contradicting the prior views that LTP and LTD occur at synapses onto distinct MSPN classes.

This latter observation appears to be tightly linked to the role striatal ACh transmission plays in striatal plasticity. In an attempt to solve the paradox that D2 receptor-dependent LTD is possible in striatal MSPNs even though all do not express postsynaptic D2 receptors, Wang et al. (2006) showed that D2 receptor-dependent LTD requires the activation of *D2 receptors on striatal TANs*. Indeed, this result nicely complements those in Bonsi et al. (2008): activation of D2 receptors on striatal TANs slows the autonomous spiking of TANs, reducing ACh release. Indeed, LTD induction is reinstated by D2 receptor antagonists and by lowering postsynaptic M₁ receptor activation. Thus, D2 receptor-dependence of LTD appears to be another manifestation of ACh-dependence. These reports cohere with those of Centonze et al. (1999), and together they suggest that a pause in ACh transmission is permissive of striatal LTD induction, whereas baseline or elevated ACh level is required for striatal LTP induction. Reinforcement learning rules based on the presumption of distinct processes operating on different classes of MSPNs will have to be replaced with more realistic learning rules that reflect these factors.

Studies of interval timing shed more light on the roles of neurotransmitters in reinforcement learning in the BG. Cortico-striatal circuits have been implicated in the timing of intervals in the seconds-to-minutes range, and dopaminergic and cholinergic drugs have been reported (Meck, 1996; Buhusi and Meck, 2005) to advance or delay the transition from low- to high-rate responding that occurs when an animal expects that action-contingent reward is imminent. For example, systemic administration of DA agonists (e.g., methamphetamine) causes an immediate, proportional leftward shift in the distribution of peak response times (i.e., promotes early responses), whereas DA antagonists cause similar rightward shifts (Meck, 1983, 1986). In contrast, systemic administration of ACh agonists (such as physostigmine) produces no immediate effect, but if continued for multiple learning sessions, causes a gradual, proportional leftward shift in the distribution of peak response times (Meck, 1983; Meck and Church, 1987), whereas ACh antagonists cause rightward shifts. However, a conspicuous difference between DAergic and cholinergic involvement in adaptive interval timing is that DAergic drug effects are compensable by further learning while on the drug. When drug administration is discontinued, a temporary rebound effect with opposite latency occurs, after which the animal once again returns to the appropriate response time by further learning (Meck, 1996). The cholinergic drug effects, however, are not compensable by learning and do not show rebound effects. Based on these observations, the DA effect has been called a *performance* or “clock speed” effect, whereas the ACh effect is interpreted as altering the *learned* response time, a “memory effect”. Nevertheless, the normative role of DA and ACh in adaptive interval timing remains to be explicated. As local interactions within the cortex-BG circuits are disclosed in all their complexity, realistic models will be vital to compute their mutual implications.

2.5. Striatal acetylcholine release depends on expected value, stimulus salience, and task-relevance

Beyond its role in striatal learning and interval timing, a further implication of the task-dependent DA-ACh cascade in the striatum (Tan and Bullock, 2008a, Section 2.4) is its bearing on striatal performance functions. The dominant response of TANs to a cue-induced burst DA signal, indicating a positive RPE, is a pause followed by a rebound reactivation. Thus TAN responses reflect expected value of cues. However, TANs also receive inputs from the thalamic centromedian and parafascicular (CM-Pf) nuclei, whose responses reflect the novelty, salience, and task-relevance of cues. One computational implication, partly explicated as a bi-conditional response surface computed by Tan and Bullock (2008a), is that it is combinations of a cue's *expected value*, *perceptual salience*, and *task-relevance* that control striatal decision making, not expected value alone. One immediate consequence of interactions among these three decision variables for striatal cholinergic signaling is that, not only phasic DA elevations, but also the gradual build-up of DA during the delay period can influence the striatal decision-

making process via cholinergic transmission, especially when the stimuli have significant salience/relevance (Figure 1). Striatal ACh transmission, in turn, exerts direct control on striatal performance functions via its effects on the MSPNs that target the output nuclei of the basal ganglia (Wang and McGinty, 1997; Alcantara et al., 2001).

It has been assumed in some computational models that DA facilitates D1-SP-MSPNs while suppressing D2-ENK-MSPNs (e.g. Gurney et al., 2001a,b; Humpries et al., 2006), and that ACh has an effect opposite to DA. However, data show a more complicated picture. ACh stabilizes the prevailing MSPN state by modulating several intrinsic currents (Howe and Surmeier, 1995; Gabel and Nisenbaum, 1999; Surmeier et al., 2005). DA, in contrast, has a state-dependent effect on MSPNs (Flores-Hernandez et al., 2002; Gruber et al., 2003): it facilitates MSPN responses when MSPNs are in a depolarized (up) state while depressing MSPNs in a hyperpolarized (down) state. Therefore, it is probable that the response of striatal MSPNs to a given corticostriatal glutamatergic input *in vivo* strongly depends on the patterning within the cascading DA-ACh signal (cf. Tan and Bullock, 2008a, see also Section 2.4) in the striatum. This contrasts to the common presumptions of the actor-critic architecture that (1) phasic DA level in the striatum has only a direct effect on striatal processing by exciting or inhibiting striatal output neurons (e.g. Frank, 2005; Frank and O'Reilly, 2006), and (2) striatal action-gating is a simple linear or sigmoidal function of expected reward value. Contrary to these ideas, emergent interactions among various neurotransmitters (glutamate, ACh, DA, GABA) in the striatum provide a much more flexible schema governing striatal performance functions, involving an interplay among at least the three aforementioned decision variables. Therefore, performance rules governing the actor component of the actor-critic system also need modifications to better reflect biological reality.

Whether any decision centers other than striatum can offer similar, to say nothing of better, sensitivity to multiple desiderata, remains to be seen. One key region to consider is the orbitofrontal cortex, which has been strongly implicated in the ability to resist framing effects in decision making (DeMartino et al., 2006), and which has been modeled recently (Draniias et al., 2008) as a key nexus in an evaluative neuraxis that includes the hypothalamus and amygdala (cf. also Frank and Claus, 2006).

2.6. Should any nuclei of the extended amygdala be regarded as parts of the striatum?

Swanson (2000) has argued for extending the concept of the BG to encompass further forebrain nuclei, notably parts of the “extended amygdala” (e.g., de Olmos and Heimer, 1999). Although some nuclei of the amygdala are far more “cortical” (because the principle neurons are glutamatergic) than “striatal”, McDonald (2003) suggested a “consensus ... that the lateral portions of the central nucleus [of the amygdala] are striatal-like (p. 13).” Notably, this region does not reciprocate its projections from cortex, its principle neurons are GABAergic MSPNs, and it connects appropriately with midbrain DA neurons. However, from a computational perspective, dis-analogies are equally important, if they imply that a generic striatal circuit model cannot be used to simulate processing in CeA (central nucleus of amygdala). Two dis-analogies may prove to be decisive. First, it is generally believed that the CeA's ACh is supplied by a everts arriving from basal forebrain (e.g. Schaafer et al., 1988), and not by intrinsic giant cholinergic neurons (TANs) like those found in “traditional” striatum. Second, although output from CeA MSPNs is potently regulated by feedforward inhibition (Paré et al., 2003), Zahm et al. (2003) found that the GABAergic parvalbumin immunoreactive (PV⁺) fast-spiking interneurons (FS-INs) characteristic of the striatum are absent from the extended amygdala. Both differences would preclude readily adapting any computational model of normal striatum to simulate information processing in lateral CeA, the “most striatal” part of the amygdala. That said, it must be admitted that even within the traditional striatum, any model must be adapted to capture important regional variations.

2.7. Basal ganglia architecture cannot be captured in feed-forward models: Toward a better understanding of perseverance, lockout protection, interrupts, switching, and resumption

Most published representations of the BG are so incomplete as to promote severe underestimations of the computational competence of the BG. That situation is being rectified by some of the microcircuit models noted above. However, the BG macrocircuit is also much different than typically depicted. First, the typical depiction aptly emphasizes that the cortico-striatal projection is not reciprocated by a striato-cortical projection. This promotes thinking of the BG as a structure dominated by a feed-forward flow along the path: cortex-striatum-pallidum-thalamus (and back to cortex). However, as briefly explored in Brown et al. (2004) (see also Srihasam et al., 2009), and as the more complete circuit in Figure 2B suggests even more emphatically, the “feedforward BG” conception has little basis. In fact, Brown et al. (2004) showed the importance of recognizing that many of the cells of origin of the cortico-striatal projection are not identical to the cells of origin of the cortico-STN projection (see also Turner and DeLong, 2000). Because of their non-identity, the former class can serve as plan representations, whereas the latter can be activated only at time of plan execution. The projection of these cells’ output back into the BG via the STN can then be understood as helping to lockout competing plans and thereby provide the selected plan enough time to execute - at least in the general case.

Another conspicuous “feedback flow” is mediated by the projection from GPe back to the striatum. This projection originates from a subset of GABAergic parvalbumin immunoreactive (PV⁺) neurons that are recipients of STN projections, and targets exclusively the PV⁺ fast-spiking interneurons (FS-INs) in the striatum (Bevan et al., 1998). Though this feedback projection has been neglected by most modelers, anatomical considerations shed some light on its functional implications. Striatal MSPNs and FSINs receive similar plan-related inputs. The prominence of collaterals between MSPNs has been taken as evidence for a winner-take-all competition between these output neurons. However, recent data challenged this assumption, showing that inhibitory communication occurs almost exclusively between different classes of MSPNs, and is not reciprocated (Venance et al., 2004; Taverna et al., 2008, see also below). Such data and other considerations inspired the proposal that feed-forward inhibition via striatal FS-INs mediates striatal competition and selection (e.g., Brown et al., 2004; Bullock and Tan, 2007). Furthermore, striatal FS-INs are coupled via gap-junctions (Koos and Tepper, 1999; Tepper et al., 2004). Such coupling can promote synchronous activity of FS-INs, yet preserves topographic organization by allowing cortico-striatal terminals with restricted distributions to nevertheless recruit FS-INs broadly, allowing robust feedforward inhibition of striatal MSPNs. Thus, it is conceivable that selection of a plan among several others is highly sensitive to the relative cortical activation levels and/or corticostriatal synaptic weights at the moment when one plan wins the competition and activates a corresponding MSPN despite feedforward inhibition via FS-INs. However, if FS-IN inhibition of the winning MSPN were to persist during the entire movement interval, the winning MSPN would remain at risk of falling below its activation threshold, especially if competing cortical plan representations remain active and are of nearly equal strength. Nevertheless, this “risk” can be reduced by channel-specific inhibitory feedback from GPe to FS-INs that are in the neighborhood of the winning MSPN: cortico-subthalamo-pallidal projections can activate cells of origin of the pallido-striatal feedback pathway, thereby inhibiting the subset of FS-INs that are recipients of selected cortical plan representations, disinhibiting corresponding MSPNs. Therefore, one possible computational role of the feedback projection from GPe to striatal FS-INs is to enable a “real-time contrast-enhancement” at the striatum while the outcome of the ongoing selection process is still unfolding.

Added to this emerging picture are the “horizontal” interactions of two classes of MSPNs. As mentioned above, the assumption of a winner-take-all (WTA) competition between D1-SP-

and D2-ENK-MSPNs has been challenged on the basis of electrophysiology (Jaeger et al., 1994) and logic (Brown et al., 2004, p. 476). Here we add that the breadth of collateral arborization falls far short of what would be needed to achieve WTA selection across broad regions of the striatum. Furthermore, recent data show that: (1) a subset of MSPNs in the striatum are coupled via gap junctions (Onn and Grace, 1994; Venance et al., 2004); (2) this electrotonic coupling is mostly confined to D2-ENK-MSPNs (Onn and Grace, 1994; Venance et al., 2004); (3) chemical (GABAergic) transmission between MSPNs is potent but unidirectional (Tunstall et al., 2002; Venance et al., 2004); and, equally important, (4) electrotonic and unidirectional chemical communication among MSPNs are mutually exclusive (Venance et al., 2004). The first implication is that D2-ENK-MSPNs do not inhibit each other. That D2-ENK-MSPNs are not mutually inhibitory complements their gap-junction coupling because it allows even a focused glutamatergic input of sufficient strength and duration to recruit them synchronously and in significant numbers. This is consistent with the “STOP” or “NO-GO” function attributed to D2-ENK-MSPNs in some models (e.g., Brown et al., 2004; Frank, 2005), whereas, had the contrary been found, i.e., had it been found that D1-SP-MSPNs were both electro-tonically coupled and not mutually inhibitory, it would have disconfirmed all recent BG models. Beyond this key conclusion, the data force a choice between two possibilities in any model. Either active D1-SP-MSPNs inhibit D2-ENK-MSPNs, or vice versa, but not both (because chemical transmission is non-reciprocated). Here it is important to recall that because of the limited arborization of MSPN feedback collaterals, feedback inhibition by D2-ENK-MSPNs would be much more powerful (than by D1-SP-MSPNs), because their electrotonic coupling would enable feedback inhibition to be much broader than that overtly implied by the limited arborization. This would nicely enhance the proposed “STOP” signal function of the indirect pathway. By contrast, D1-SP-MSPN feedback inhibition would remain highly focused, would not support a WTA property, would not sharpen contrast among competing direct pathway MSPNs, but would interfere with the STOP-signal function by inhibiting nearby D2-ENK-MSPNs. From a global functional perspective, it therefore seems much more likely that the observed unidirectional GABAergic transmission between MSPNs (Venance et al., 2004) runs from D2-ENK-MSPNs to D1-SP-MSPNs. In fact, it has recently been shown that feedback inhibitory communication directed from D2-ENK-MSPNs to D1-SP-MSPNs is much more prominent than the alternatives (Taverna et al., 2008).

The difference in neuropeptide co-release by different classes of MSPNs (substance-P vs. enkephalin) begets a further asymmetry. Cortico-striatal terminals in the striatum express presynaptic neurokinin-1 receptors (primary target of substance-P in primates and humans; Regoli et al., 1994), and it has recently been shown that endogenous SP released by D1-SP-MSPNs may enhance presynaptic glutamate release, thereby facilitating postsynaptic responses in neighboring MSPNs (Blomeley et al., 2009). It should be noted that only D1-SP-MSPNs can partake in such a feedback *excitation* by virtue of somatically co-releasing SP with GABA. Through this feedback excitation, a subset of MSPNs, once selected, can recruit neighboring MSPNs (presumably belonging to the same functional channel), perhaps to further enhance the contrast between selected and competing plan representations.

The data summarized above clearly suggest that local interactions within the BG, particularly in the striatum, are not only more complicated than assumed in most models of the BG, but also can offer substantial computational abilities that most current models of the BG lack. Nevertheless, human intuition is generally insufficient to predict ramifications of such complex interactions, and computational models that reflect these microcircuit specializations can reveal their implications for BG functions. Recently, Bullock and Tan (2007) used the more complete macro-circuit in Figures 2B and 2C as a basis for exploring how the BG circuit enables a more complete set of fundamental abilities that serve as a basis for intelligent cognitive control of behavioral scheduling, sequencing, and interleaving, including the ability to interrupt, switch,

and resume planned behavior. Demonstration of such abilities of the BG, with models that take a more complete macro- and microcircuit of the BG into account than traditional TD- and RPE-based actor-critic architecture, in turn, opens up exciting avenues for integrating the BG into computational models of higher cognitive and behavioral functions.

2.8. Basal ganglia architecture is pivotal for advanced human abilities, including speech

To some who considered language as the most “neo” of neo-cortical functions, it came as a shock when early fMRI studies (e.g., Ullman et al., 1997) strongly implicated the BG in such prototypical linguistic functions as control of regular past-tense production (e.g., postfixing –*ed* in English). The BG have since also been implicated in arithmetic rule application (Teichmann et al., 2008). Despite such data clues, few computational models of speech, linguistic or arithmetic rule application make integral use of either the cortex-thalamus-BG macrocircuit or BG microcircuits. Steps to rectify this shortfall have recently been taken by several research groups. For example, Bohland et al. (2009) are exploring the hypothesis that the BG are critical for ensuring that speech acts satisfy the multiple types of constraints that must be met by linguistic productions if they are to succeed as conventional communications.

Consider the difference between computational models of de-contextualized sequence production and meaning-communicative sequence production, which obeys learned rules. For example, chronometric (latency) patterns of anticipatory or deferral errors in non-linguistic sequence production (Farrell and Lewandowsky, 2004) as well as electrophysiological recordings (Averbeck et al., 2002; Rhodes et al., 2004) have strongly supported a class of sequence control models, called *competitive queuing* (CQ) models (Averbeck et al., 2002; Bullock, 2004a,b; Rhodes et al., 2004; Ivey et al., 2008), that are defined by two assumptions (cf. Grossberg, 1978): (1) the sequential order relation among plan-representations of forthcoming acts is represented by an analog gradient (“primacy gradient”) of activation levels established over the plan representations in a WM (working memory), and (2) once a plan representation is chosen for enactment, its representation is deleted from the planning WM, and thus eliminated from the competition (among the surviving representations) that determines which plan to perform next. However, it has not been clear whether CQ models should, or how they could, be extended to explain linguistic sequences controls. For example, although many linguistic sequencing errors are exchange errors, as predicted by CQ theory, elementary CQ theory does not explain why linguistic exchange errors respect linguistic class. The connectionist language production model of Ward (1994), which utilized concepts from construction grammar (e.g., Goldberg, 2006), explored one way that a CQ process could be extended to ensure that next-word choices simultaneously obeyed semantic and syntactic constraints. However, Ward (1994) offered no interpretation of model components in terms of identified brain circuits. Although it did not include BG microcircuits, the computational model of language processing in Dominey et al. (2006) also adopted ideas from construction grammar, and proposed that the cortical-striatal projection mediates retrieval of form-to-meaning mappings. The computational model of Bohland et al. (2009) incorporated macro- and some micro-circuit details from the Brown et al. (2004) model of fronto-BG function to illustrate one way to use BG circuitry to offer a CQ-consistent explanation for multi-syllabic speech production. In this model, exchange errors are appropriately class-constrained, a consequence of the model's ability to ensure that next-sound choices obey both phonemic and syllabic constraints. The model's mapping of computations to neurobiological circuits enabled it to pinpoint candidate neural bases of speech stuttering errors (Civier et al., 2009), and similar future models should be able to use BG computations to ensure the simultaneous satisfaction of multiple types of constraints in language processing, e.g., to achieve the integrative-well-formedness checks, based on semantic, syntactic, and pragmatic constraints, that were recently attributed to the BG by Bornkessel and Schlesewsky (2006).

In such approaches, the BG are seen to offer computational resources to ensure that decisions are not finalized unless and until multiple types of preconditions for success are simultaneously satisfied. This way of thinking about the BG dates back (at least) to Passingham (1987), who argued that the multiple types of information that need to be considered for good decisions often are not brought together in any single cortical region, but are brought together in compact regions of the striatum. Consistently, Brown et al. (2004) showed how cortico-striatal convergence patterns and intrinsic BG circuitry can ensure that plans are withheld from enactment until distinct types of representations, computed in multiple cortical areas, become simultaneously active, and thus coherently support performance of an associated plan.

Much work remains to fully understand how BG microcircuits support such computations, and there yet exists no consensus that the BG are obligatorily involved in language processing. A barrier to consensus is that many researchers interpret their findings with respect to a mistaken or a highly incomplete mental model of pertinent circuitry. For example, Wahl et al. (2008) used a *lack* of task-related signal variance in their recordings from human GPi and STN (during a language comprehension task) to argue that “syntactic and semantic language analysis is primarily realized within cortico-thalamic networks, whereas a cohesive basal ganglia network is not involved in these essential operations of language analysis.” A close examination of their argument reveals a number of problems. First, they concluded that the task-related signal variance that they observed in the VIM thalamus could not have been BG-dependent, because of the absence of GPi modulation. This is mistaken, because it presupposes that the only trans-BG path to the thalamus runs through the GPi. This ignores both the well-known SNr projection to thalamus, as well as another path that is little-known but even more pertinent here. As shown in Figure 2B, there is also a trans-BG path that runs from cortex to D2-ENK-MSPNs (in striatum), to the GPe, to the reticular nucleus of the thalamus, and finally to specific nuclei of the thalamus, such as VIM. Second, although they measured activity only in VIM, their interpretation, based on a review of subcortical aphasia by Nadeau and Crosson (1997), emphasizes linguistic roles for the CM and pulvinar nuclei of the thalamus. Supposing that these thalamic nuclei (not recorded in their experiments) are implicated in linguistic computations, it is hard to understand how the BG are not also strongly implicated, for two reasons: the pulvinar has maintained strong projections to the striatum since before the cerebral cortex evolved, and we earlier noted that the CM and Pf nuclei of the thalamus are potent sources of inputs to cholinergic TANs located in, respectively, the putamen and the caudate nuclei of the striatum. Moreover, there is accumulating evidence that jointly implicates the CM and BG in speech control disorders, notably stuttering (Alm, 2004; Civier et al., 2009).

3. Conclusion: Accurate Microcircuit Models are Key for Understanding Individual Differences in the Neurophysiology of Cognitive Functions

Many macro- and microcircuit specializations in the BG-frontal cortex circuits go beyond what was predicted by traditional TD models, RPE theory, and actor-critic architectures. For the traditional models to capture the computational roles of the BG implied by such specializations, both commonly used reinforcement learning rules, and performance-related rules, must be updated. Most notably, these specializations imply a far-reaching computational role for the BG-frontal cortex circuits not only in reinforcement learning but also in robust evaluation, selection, and execution of actions associated with different environmental contingencies. Striatal computations need to heed the diverse types of preconditions that must be met before a planned act can be expected to succeed. An exciting application of comprehensive models of BG-frontal cortex circuits, including key microcircuits, will be the growing ability to assess/diagnose neural bases of individual differences (e.g., Frank, 2005) and pathologies, and then to use individualized computer brain models to predict individual differences in response to therapeutic regimes, such as those based on pharmacological measures or implanted neurostimulators.

Acknowledgments

This work was supported in part by the U.S. National Science Foundation under Science of Learning Center Grant SBE-354378 and in part by NIH Grant R01DC007683.

References

- Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. *Trends Neurosci* 1989;12:366. [PubMed: 2479133]
- Alcantara AA, Mrzljak L, Jakab RL, Levey AI, Hersch SM, Goldman-Rakic PS. Muscarinic m1 and m2 receptor proteins in local circuit and projection neurons of the primate striatum: Anatomical evidence for cholinergic modulation of glutamatergic prefronto-striatal pathways. *J Comp Neurol* 2001;434:445–460. [PubMed: 11343292]
- Alm PA. Stuttering and the basal ganglia: a critical review of possible relations. *J Commun Disord* 2004;37:325–369. [PubMed: 15159193]
- Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP. Parallel processing of serial movements in prefrontal cortex. *Proc Natl Acad Sci USA* 2002;99:13172–13177. [PubMed: 12242330]
- Bevan MD, Booth PA, Eaton SA, Bolam JP. Selective innervation of neostriatal interneurons by a subclass of neuron in the globus pallidus of the rat. *J Neurosci* 1998;18:9438–9452. [PubMed: 9801382]
- Blomeley CP, Kehoe LA, Bracci E. Substance P mediates excitatory interactions between striatal projection neurons. *J Neurosci* 2009;29:4953–4963. [PubMed: 19369564]
- Bohland JW, Bullock D, Guenther FH. Neural representations and mechanisms for the performance of simple speech sequences. *J Cogn Neurosci*. 2009 in press
- Bonsi P, Martella G, Cuomo D, Platania P, Sciamanna G, Bernardi G, Wess J, Pisani A. Loss of muscarinic autoreceptor function impairs long-term depression but not long-term potentiation in the striatum. *J Neurosci* 2008;28:6258. [PubMed: 18550768]
- Bornkessel I, Schlesewsky M. The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychol Rev* 2006;113:787–821. [PubMed: 17014303]
- Brown JW, Bullock D, Grossberg S. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Netw* 2004;17:471–510. [PubMed: 15109680]
- Buhusi CV, Meck WH. What makes us tick? functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 2005;6:755–765. [PubMed: 16163383]
- Bullock D. Adaptive neural models of queuing and timing in fluent action. *Trends Cogn Sci* 2004a;8:426–433. [PubMed: 15350244]
- Bullock D. From parallel sequence representations to calligraphic control: a conspiracy of neural circuits. *Motor Control* 2004b;8:371–391. [PubMed: 15585895]
- Bullock D, Tan CO. Role of basal ganglia in decision sequences: A computational model of dopamine and acetylcholine regulation of action selection, interruption, resumptions, and switching. *Soc Neurosci Abs*. 2007page S514
- Bullock D, Tan CO. Computational implications of microcircuit specializations in forebrain circuits for motivated action selection. *IEEE Proc Int Joint Conf Neur Netw*. 2009
- Calabresi P, Maj R, Pisani A, Mercuri NB, Bernardi G. Long-term synaptic depression in the striatum: physiological and pharmacological characterization. *J Neurosci* 1992a;12:4224–4233. [PubMed: 1359031]
- Calabresi P, Pisani A, Mercuri NB, Bernardi G. Long-term potentiation in the striatum is unmasked by removing the voltage-dependent magnesium block of NMDA receptor channels. *Eur J Neurosci* 1992b;4:929–935. [PubMed: 12106428]
- Centonze D, Gubellini P, Bernardi G, Calabresi P. Permissive role of interneurons in cortical synaptic plasticity. *Brain Res Rev* 1999;31:1–5. [PubMed: 10611492]
- Civier, A.; Bullock, D.; Max, L.; Guenther, F. Simulating neural impairments to syllable-level command generation in stuttering.. Sixth World Congress on Fluency Disorders; Rio de Janeiro, Brazil. 5 – 8 August; 2009.

- de Olmos JS, Heimer L. The concepts of the ventral striatopallidal system and extended amygdala. *Ann N Y Acad Sci* 1999;877:1–32. [PubMed: 10415640]
- DeMartino B, Kumaran D, Seymour B, Dolan RJ. Frames, biases, and rational decision-making in the human brain. *Science* 2006;313:684–687. [PubMed: 16888142]
- Dominey PF, Hoen M, Inui T. A neurolinguistic model of grammatical construction processing. *J Cogn Neurosci* 2006;18:2088–2107. [PubMed: 17129193]
- Dranias MR, Grossberg S, Bullock D. Dopaminergic and non-dopaminergic value systems in conditioning and outcome-specific revaluation. *Brain Res* 2008;1238:239–287. [PubMed: 18674518]
- Farrell S, Lewandowsky S. Modeling transposition latencies: Constraints for theories of serial order memory. *J Mem Lang* 2004;51:115–135.
- Fino E, Glowinski J, Venance L. Bidirectional activity-dependent plasticity at corticostriatal synapses. *J Neurosci* 2005;25:11279–11287. [PubMed: 16339023]
- Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 2003;299:1898–1902. [PubMed: 12649484]
- Fiorillo CD, Tobler PN, Schultz W. Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behav Brain Funct* 2005;1:7. [PubMed: 15958162]
- Flores-Hernandez J, Cepeda C, Hernandez-Echeagaray E, Calvert CR, Jokel ES, Fienberg AA, Greengard P, Levine MS. Dopamine enhancement of NMDA currents in dissociated medium-sized striatal neurons: Role of D1 receptors and DARPP-32. *J Neurophysiol* 2002;88:3010–3020.
- Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. *J Cogn Neurosci* 2005;17:51–72. [PubMed: 15701239]
- Frank MJ, Claus ED. Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 2006;113:300–326. [PubMed: 16637763]
- Frank MJ, O'Reilly RC. A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* 2006;120:497–517. [PubMed: 16768602]
- Gabel LA, Nisenbaum ES. Muscarinic receptors differentially modulate the persistent potassium currents in striatal spiny projection neurons. *J Neurophysiol* 1999;81:1418–1423. [PubMed: 10085367]
- Goldberg, AE. *Constructions at work: The nature of generalization in language*. Oxford; New York: 2006.
- Grossberg, S. A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans.. In: Rosen, R.; Snell, F., editors. *Progress in Theoretical Biology*. Vol. 5. Academic Press; New York: 1978. p. 233-374.
- Gruber AJ, Solla SA, Surmeier DJ, Houk JC. Modulation of striatal single units by expected reward: A spiny neurons model displaying dopamine-induced bistability. *J Neurophysiol* 2003;90:1095–1114. [PubMed: 12649314]
- Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. i. a new functional anatomy. *Biol Cybern* 2001a;84:401–410. [PubMed: 11417052]
- Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. ii. analysis and simulation of behavior. *Biol Cybern* 2001b;84:411–423. [PubMed: 11417053]
- Hasselmo ME. The role of acetylcholine in learning and memory. *Curr Opin. Neurobiol* 2006;16:710–715. [PubMed: 17011181]
- Houk, J.; Adams, JL.; Barto, AG. *A Model of How the Basal Ganglia Generate and Use Signals That Predict Reinforcement*. MIT Press; Cambridge, MA: 1995. p. 249-270.
- Howe AR, Surmeier DJ. Muscarinic receptors modulate N-, P-, and L-type Ca²⁺ currents in rat striatal neurons through parallel pathways. *J Neurosci* 1995;15:458–469. [PubMed: 7823150]
- Humphries MD, Stewart RD, Gurney KN. A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *J Neurosci* 2006;26:12921–12942. [PubMed: 17167083]
- Ivey, R.; Bullock, D.; Grossberg, S. A neuromorphic model of spatial lookahead planning.. *Proceedings of the 12th International Conference on Cognitive and Neural Systems*; Boston, MA. 2008.

- Jaeger D, Kita H, Wilson CJ. Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. *J Neurophysiol* 1994;72:2555–2558. [PubMed: 7884483]
- Kakade S, Dayan P. Dopamine: Generalization and bonuses. *Neur Netw* 2002;15:549–559.
- Kelley AE, Cadore M, Stinus L, Le MM. Neurotensin, substance P, neurokinin-alpha, and enkephalin: Injection into ventral tegmental area in the rat produces differential effects on operant responding. *Psychopharmacology (Berl)* 1989;97:243–252. [PubMed: 2471221]
- Kerr JND, Wickens JR. Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. *J Neurophysiol* 2001;85:117–124. [PubMed: 11152712]
- Kilgard MP, Merzenich MM. Cortical map reorganization enabled by nucleus basalis activity. *Science* 1998;279:1714–1718. [PubMed: 9497289]
- Koos T, Tepper JM. Inhibitory control of neostriatal projection neurons by GABAergic interneurons. *Nat Neurosci* 1999;2:467–472. [PubMed: 10321252]
- Laurent PA. The emergence of saliency and novelty responses from reinforcement learning principles. *Neur Netw* 2008;21:1493–1499.
- McDonald AJ. Is there an amygdala and how far does it extend? An anatomical perspective. *Ann N Y Acad Sci* 2003;985:1–21. [PubMed: 12724144]
- Meck WH. Selective adjustment of the speed of internal clock and memory processes. *J Exp Psychol Anim Behav Proc* 1983;9:171.
- Meck WH. A utility for the dopamine D2 receptor predicts neuroleptic potency in decreasing the speed of an internal clock. *Pharm Biochem Behav* 1986;25:1185.
- Meck WH. Neuropharmacology of timing and time perception. *Cogn Brain Res* 1996;3:227–242.
- Meck WH, Church RM. Cholinergic modulation of the content of temporal memory. *Behav Neurosci* 1987;101:457–464. [PubMed: 2820435]
- Nadeau SE, Crosson B. Subcortical aphasia. *Brain Lang* 1997;58:355–402. [PubMed: 9222518]
- Niv Y, Du MO, Dayan P. Dopamine, uncertainty and td learning. *Behav Brain Funct* 2005;1:6. [PubMed: 15953384]
- Onn S-P, Grace AA. Dye coupling between rat striatal neurons recorded in vivo: compartmental organization and modulation by dopamine. *J Neurophysiol* 1994;71:1917–1934. [PubMed: 8064357]
- Paré D, Royer S, Smith Y, Lang EJ. Contextual inhibitory gating of impulse traffic in the intra-amygdaloid network. *Ann N Y Acad Sci* 2003;985:78–91. [PubMed: 12724150]
- Passingham, RE. From where does the motor cortex get its instructions?. In: Wise, SP., editor. *Higher brain functions*. Wiley; New York: 1987. p. 67-97.
- Pawlak V, Kerr JN. Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J Neurosci* 2008;28:2435–2446. [PubMed: 18322089]
- Placenza FM, Fletcher PJ, Rotzinger S, Vaccarino FJ. Infusion of the substance P analogue, DiMe-C7, into the ventral tegmental area induces reinstatement of cocaine-seeking behaviour in rats. *Psychopharmacology (Berl)* 2004;177:111–120. [PubMed: 15167979]
- Redgrave P, Prescott TJ, Gurney K. The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* 1999a;89:1009–1023. [PubMed: 10362291]
- Redgrave P, Prescott TJ, Gurney K. Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 1999b;22:146–151. [PubMed: 10203849]
- Regoli D, Boudon A, Fauchere J-L. Receptors and antagonists for substance P and related peptides. *Pharmacol Rev* 1994;45:551–599. [PubMed: 7534932]
- Rhodes BJ, Bullock D, Verwey WB, Averbeck BB, Page MP. Learning and production of movement sequences: Behavioral, neurophysiological, and modeling perspectives. *Hum Mov Sci* 2004;23:699–746. [PubMed: 15589629]
- Schäfer MK, Eiden LE, Weihe E. Cholinergic neurons and terminal fields revealed by immunohistochemistry for the vesicular acetylcholine transporter. I. Central nervous system. *Neuroscience* 1988;1998:331–359.
- Schotanus SM, Chergui K. Dopamine D1 receptors and group I metabotropic glutamate receptors contribute to the induction of long-term potentiation in the nucleus accumbens. *Neuropharmacology* 2008;54:837–844. [PubMed: 18272187]

- Sjöström PJ, Nelson SB. Spike timing, calcium signals, and synaptic plasticity. *Curr Opin Neurobiol* 2002;12:305–314. [PubMed: 12049938]
- Srihasam K, Bullock D, Grossberg S. Target selection by frontal cortex during coordinated saccadic and smooth pursuit eye movements. *J Cogn Neurosci* 2009;21in press
- Surmeier DJ, Mercer JN, Chan CS. Autonomous pacemakers in the basal ganglia: who needs excitatory synapses anyway? *Curr Opin Neurobiol* 2005;15:312–318. [PubMed: 15916893]
- Sutton RS, Barto AG. Toward a modern theory of adaptive networks: Expectation and prediction. *Psychol Rev* 1981;88:135–170. [PubMed: 7291377]
- Swanson LW. Cerebral hemisphere regulation of motivated behavior. *Brain Res* 2000;886:113–164. [PubMed: 11119693]
- Tan CO, Anderson E, Dranias M, Bullock D. Can the apparent adaptation of dopamine neurons' mismatch sensitivities be reconciled with their computation of reward prediction errors? *Neurosci Lett* 2008;438:14–16. [PubMed: 18482798]
- Tan CO, Bullock D. A dopamine-acetylcholine cascade: Simulating learned and lesion-induced behavior of striatal cholinergic interneurons. *J Neurophysiol* 2008a;100:2409–2421. [PubMed: 18715897]
- Tan CO, Bullock D. A local circuit model of learned striatal and dopamine cell responses under probabilistic schedules of reward. *J Neurosci* 2008b;28:10062–10074. [PubMed: 18829964]
- Tan CO, Bullock D. Neuropeptide co-release with GABA may explain functional non-monotonic uncertainty responses in dopamine neurons. *Neurosci Lett* 2008c;430:218–223. [PubMed: 18053647]
- Taverna S, Ilijic E, Surmeier DJ. Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of parkinson's disease. *J Neurosci* 2008;28:5504–5512. [PubMed: 18495884]
- Teichmann M, Guara V, Démonet JF, Supiot F, Delliaux M, Verny C, Renou P, Remy P, Bachoud-Lévy AC. Language processing within the striatum: evidence from a PET correlation study in Huntington's disease. *Brain* 2008;131:1046–1056. [PubMed: 18334537]
- Tepper JM, Koos T, Wilson CJ. GABAergic microcircuits in the neostriatum. *Trends Neurosci* 2004;27:662–669. [PubMed: 15474166]
- Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science* 2005;307:1642–1645. [PubMed: 15761155]
- Tunstall MJ, Oorschot DE, Kean A, Wickens JR. Inhibitory interactions between spiny projection neurons in the rat striatum. *J Neurophysiol* 2002;88:1263–1269. [PubMed: 12205147]
- Turner RS, DeLong MR. Corticostriatal activity in primary motor cortex of the macaque. *J Neurosci* 2000;20:7096–7108. [PubMed: 10995857]
- Ullman MT, Corkin S, Coppola M, Hickok G, Growdon JH, Koroshetz WJ, Pinker S. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *J Cogn Neurosci* 1997;9:266–276.
- Ungless MA, Magill PJ, Bolam JP. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 2004;303:2040–2042. [PubMed: 15044807]
- Venance L, Glowinski J, Giaume C. Electrical and chemical transmission between striatal GABAergic output neurones in rat brain slices. *J Physiol* 2004;559:215–230. [PubMed: 15235091]
- Wahl M, Marzinzik F, Friederici AD, Hahne A, Kupsch A, Schneider GH, Saddy D, Curio G, Klostermann F. The human thalamus processes syntactic and semantic language violations. *Neuron* 2008;59:695–707. [PubMed: 18786354]
- Wang JQ, McGinty JF. The full D1 dopamine receptor agonist SKF-82958 induces neuropeptide mRNA in the normosensitive striatum of rats: Regulation of D1/D2 interactions by muscarinic receptors. *J Pharmacol Exp Ther* 1997;281:972–982. [PubMed: 9152408]
- Wang Z, Kai L, Day M, Ronesi J, Yin HH, Ding J, Tkatch T, Lovinger DM, Surmeier DJ. Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* 2006;50:443–452. [PubMed: 16675398]
- Ward, N. *A connectionist language generator*. Ablex Publishing; Norwood, NJ: 1994.
- Zahm DS, Grosu S, Irving JC, Williams EA. Discrimination of striatopallidum and extended amygdala in the rat: a role for parvalbumin immunoreactive neurons? *Brain Res* 2003;978:141–154. [PubMed: 12834908]

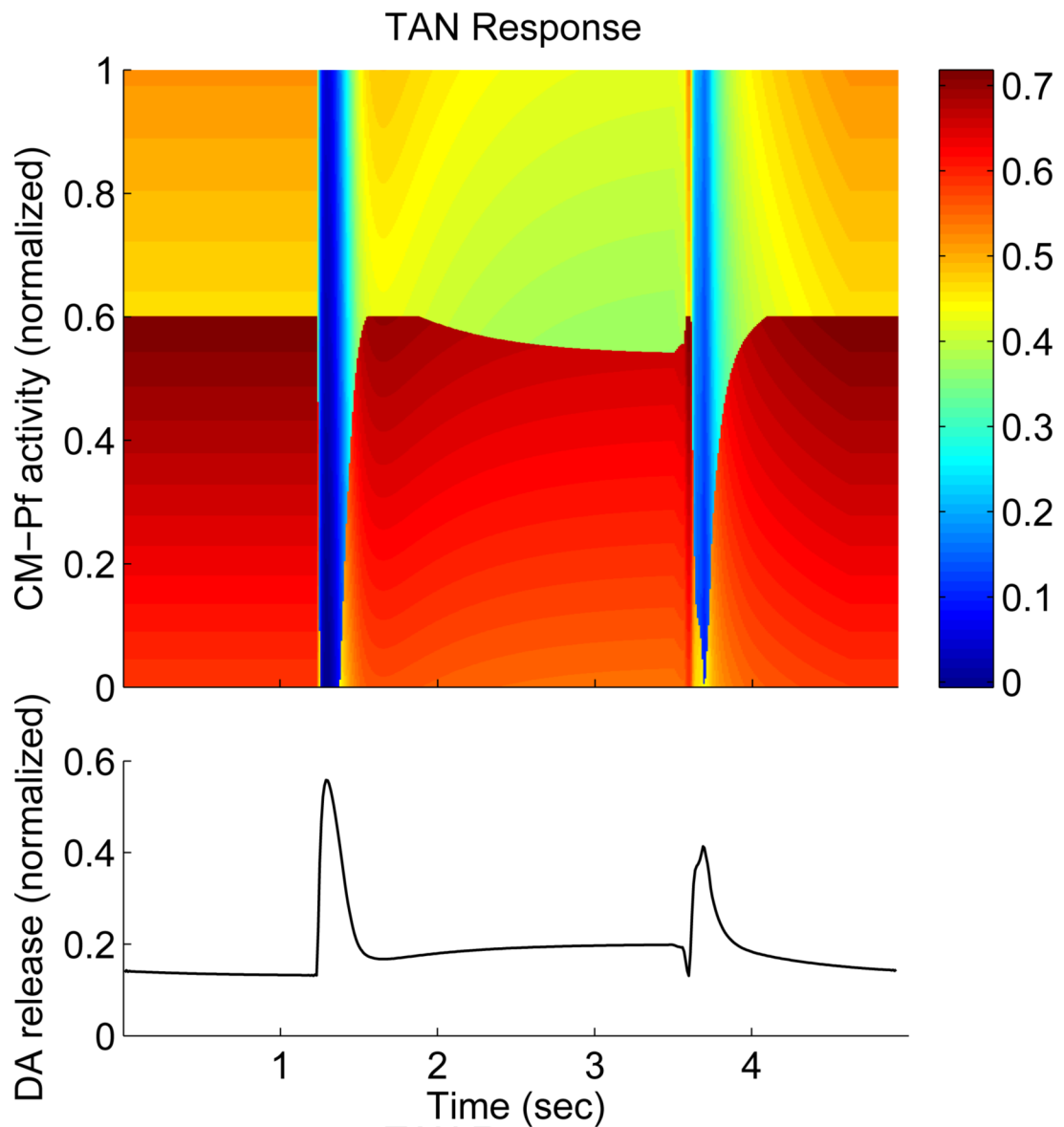


Figure 1.

The tight coupling between striatal DAergic, cholinergic, and thalamo-striatal glutamatergic transmission. The upper panel's color values show the normalized cholinergic interneuron (TAN) responses (baseline firing rate 0.5), if allowed to reach equilibrium (Tan and Bullock, 2008a), as a function both of phasic and sustained dopamine signal inputs (shown in the lower panel) that reflect reward expectations, and of thalamo-striatal glutamatergic inputs (from CM-Pf nuclei, shown on the ordinate) that reflect stimulus salience. The lower panel shows the (normalized) learned DA response to a CS-then-reward sequence, after training with a $p = 0.5$ conditional probability schedule during delay conditioning with a 2 second-long CS (Tan and Bullock, 2008b). The phasic DA burst in response to CS onset ($t = 1.2$ sec) reflects the expected value of reward; that in response to primary reward ($t = 3.2$ sec) reflects the probabilistic schedule; and the gradual increase during the delay period reflects the uncertainty in the expected outcome (Tobler et al., 2005; Tan and Bullock, 2008b; Fiorillo et al., 2003, 2005). Although the TAN responses (upper panel) also reflect such task-dependent variables, TAN responses also depend upon stimulus salience.

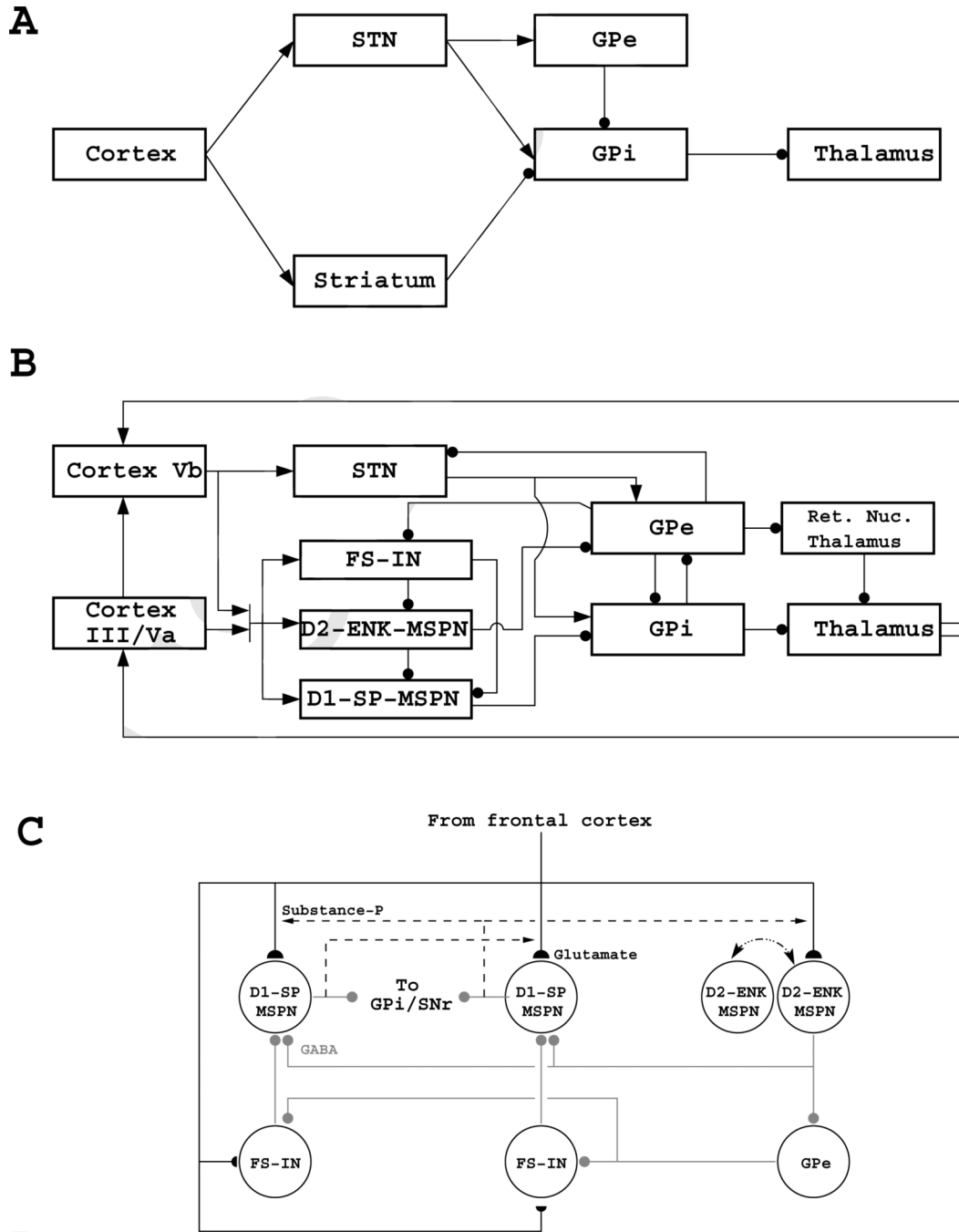


Figure 2. The basal ganglia (BG) system is not correctly represented as a feedforward structure. Panel A depicts a common view of BG information processing. Panel B adds further details to clarify some of the multiple kinds of feedback that need to be considered in a realistic model. Panel C shows the microcircuit-level interactions between medium spiny projection neurons and fast-spiking interneurons in the striatum. Black arrows denote excitatory connections, black semi-circles denote adaptive synaptic connections, black circles (and gray lines/circles in Panel C) denote inhibitory connections. Key: Vb, layer Vb of frontal cortex; III/Va, layers III and Va of frontal cortex; STN, subthalamic nucleus; FS-IN, fast-spiking GABAergic interneuron of the striatum; D2-ENK-MSPN, a dopamine D2-type receptor-expressing, enkephalin co-

releasing, GABAergic medium spiny projection neuron of the striatum; D1-SP-MSPN, a dopamine D1-type receptor-expressing, substance P co-releasing, GABAergic medium spiny projection neuron of the striatum; GPe, external (lateral) globus pallidus; GPi, internal (medial) globus pallidus; Ret. Nuc.: reticular nucleus of thalamus.