Chapter 8
Linking Across Levels of Computation in Model-Based Cognitive Neuroscience

Michael J. Frank

Abstract Computational approaches to cognitive neuroscience encompass multiple levels of analysis, from detailed biophysical models of neural activity to abstract algorithmic or normative models of cognition, with several levels in between. Despite often strong opinions on the 'right' level of modeling, there is no single panacea: attempts to link biological with higher level cognitive processes require a multitude of approaches. Here I argue that these disparate approaches should not be viewed as competitive, nor should they be accessible to only other researchers already endorsing the particular level of modeling. Rather, insights gained from one level of modeling should inform modeling endeavors at the level above and below it. One way to achieve this synergism is to link levels of modeling by quantitatively fitting the behavioral outputs of detailed mechanistic models with higher level descriptions. If the fits are reasonable (e.g., similar to those achieved when applying high level models to human behavior), one can then derive plausible links between mechanism and computation. Model-based cognitive neuroscience approaches can then be employed to manipulate or measure neural function motivated by the candidate mechanisms, and to test whether these are related to high level model parameters. I describe several examples of this approach in the domain of reward-based learning, cognitive control, and decision making and show how neural and algorithmic models have each informed or refined the other.

8.1 Introduction

Cognitive neuroscience is inherently interested in linking levels of analysis, from biological mechanism to cognitive and behavioral phenomena. But there are not just two levels, rather, a continuum of many. One can consider the implications of particular ion channel conductances and receptors, the morphological structure of individual neurons, the translation of mRNA, intracellular molecular signaling cascades involved in synaptic plasticity, and so forth. At the cognitive level, the

M. J. Frank (✉)
Cognitive, Linguistic & Psychological Sciences, Brown Institute for Brain Science, Providence, USA
e-mail: Michael_Frank@Brown.edu

B. U. Forstmann, E.-J. Wagenmakers (eds.), *Model-Based Cognitive Neuroscience,*
DOI 10.1007/978-1-4939-2236-9_8

field typically discusses constructs such as working memory, executive control, reinforcement learning, episodic memory, to name a few. In between these levels reside architectures of neural systems, such as the frontal cortex, parietal cortex, hippocampus and basal ganglia, the interactions among all of these systems, and their modulations by neurotransmitters in response to relevant task events. Computational models greatly facilitate the linking of levels of analysis, because they force one to be explicit about their assumptions, to provide a unifying coherent framework, and to specify the computational objectives of any given cognitive problem which provide constraints on interpreting the underlying mechanisms.

Nevertheless, the question remains of which level of modeling to use. The field of computational neuroscience for example typically considers how low level mechanisms can give rise to higher level "behaviors", but where behavior here is defined in terms of the changes in membrane potentials of individual neurons or even compartments within neurons, or in terms of synchrony of neural firing across populations of cells. The field of computational cognitive science, on the other hand, considers how behavioral phenomena might be interpreted as optimizing some computational goal, like minimizing effort costs, maximizing expected future reward, or optimally trading off uncertainty about multiple sources of perceptual and cognitive information to make inferences about causal structure. In between, computational cognitive neuroscience considers how mechanisms within neural systems can solve tradeoffs, for example between pattern separation and pattern completion in hippocampal networks [57], between updating and maintenance of working memory in prefrontal cortex [11, 35], or between speed and accuracy in perceptual decision making as a function of connectivity between cortex and basal ganglia [9, 53]. Even with this limited number of examples however, models took multiple levels of description, from those using detailed spiking neurons to higher level computations capturing reaction time distributions, where latent estimated parameters are correlated with neural measures extracted from functional imaging. In general, theorists and experimentalists are happy to "live" at one level of analysis which intuitively has greatest aesthetic, even though there is large variation in the appreciation of what constitutes the "right" level. Although there is a rich literature in mathematical psychology on how to select the most parsimonious model that best accounts for data without overfitting, this issue primarily pertains to applications in which one quantitatively fits data, and not to the endeavor of constructing a generative model. For example, if given only error rates and reaction time distributions, a mathematical psychologist will select a minimalist model that can best account for these data with a few free parameters, but this model will not include the internal dynamics of neural activities. Some researchers also perform model selection together with functional imaging to select the best model that accounts for correlations between brain areas and psychological parameters (e.g., [48]), but even here the neural data are relatively sparse and the observations do not include access to internal circuitry dynamics, neurotransmitters, etc—even though everyone appreciates that those dynamics drive the observed measurements.

The reason everyone appreciates this claim is that the expert cognitive neuroscientist has amassed an informative prior on valid models based on a large body

of research that spans multiple methods, species, analysis tools etc. Nevertheless, it is simply not feasible to apply these informative priors into a quantitative model selection process given much more sparse data (e.g. given BOLD fMRI data and behavioral responses one would not be advised to try to identify the latent parameters of a detailed neuronal model that includes receptor affinities, membrane potential time constants, etc).

In this chapter, I advocate an alternative, multi-level modeling strategy to address this issue. This strategy involves critical integration of the two levels to derive predictions for experiments and to perform quantitative fits to data. In one prong, theoreticians can create detailed neural models of interacting brain areas, neurotransmitters, etc, constrained by a variety of observations. These models attempt to specify interactions among multiple neural mechanisms and can show how the network dynamics recapitulate those observed in electrophysiological data, and how perturbation of those dynamics (by altering neurotransmitter levels of receptor affinities) can lead to observable changes in behavior that qualitatively match those reported in the literature. In a second prong, the modeler can construct a higher level computational model motivated by the mathematical psychological literature which summarizes the basic cognitive mechanism. Examples of this level include simple reinforcement learning models from the machine learning literature (e.g., Q learning; [73]), or sequential sampling models of decision making such as the drift diffusion model (see [60], for a review). These models should be constructed such that the parameters are identifiable, meaning that if one generates fake data from the model they should be able to reliably recover the generative parameters and to differentiate between changes that would be due to alterations in one parameter separately from other parameters. Ideally, the experimental task design will be informed by this exercise prior to conducting the experiment, so that the task can provide conditions that would be more diagnostic of differences in underlying parameters. The identifiability of a model is a property of not only of the model itself but also the different task conditions. To see this, consider a model in which task difficulty of one sort or another is thought to selectively impact a given model parameter. One might include two or more different difficulty levels, but the degree to which these influence behavior, and thus lead to observable differences that can be captured by the relevant model parameter, often interact with the other task and model parameters.

Given an identifiable model and appropriate task, the next step is to link the levels of modeling to each other. Here, one generates data from the detailed neural model exposed to the task such that it produces data at the same level as one would obtain from a given experiment—error rates and RT distributions for example, or perhaps also some summary statistic of neural activity in a given simulated brain area in one condition vs. another as one might obtain with fMRI or EEG. Then, these outputs are treated just as one does when fitting human (or other animal) data: assuming they were generated by the higher level model (or several candidate higher level models). Model selection and parameter optimization proceeds exactly as it would when fitting these models to actual human data. The purpose of this exercise is to determine which of the higher level models best summarizes the effective computations of the detailed model. Further, one can perform systematic parameter manipulations in the neural

model to determine whether these have observable selective effects on the higher level parameter estimates. If the fit is reasonable (for example if the measures of model fit are similar to those obtained by fitting the higher level model to human data), this process can lead to a higher level computational description of neural circuit function not directly afforded by the detailed neural dynamic simulations. (This approach is complementary to that taken by Bogacz and colleagues, who have used a single level of computation but made observations about how distinct aspects of the computational process can be mapped onto distinct nuclei within the cortico-basal ganglia network [7]. In Bogacz's work the computations afforded by this circuitry is identical to that of the optimal Bayesian model of decision making. In our models described below, these computations emerge from nonlinear neural dynamics and the mapping is approximate, not exact, and hence allow for potential refinements in the higher level description that best characterizes human cognition and behavior.)

When successful, this exercise can also motivate experiments in which the same biological manipulation is performed on actual participants (e.g. a medication manipulation, brain stimulation, pseudoexperimental manipulation via genetics). Indeed, by linking across levels of computation one derives precise, falsifiable predictions about which of the higher level observable and identifiable parameters will be affected, and in which direction, by the manipulation. It can also provide informative constraints on interpreting how component processes are altered as a function of mental illness [54]. Of course, one may derive these predictions intuitively based on their own understanding of neural mechanisms, but there are various instances in which explicit simulations with more detailed models can lend insight into interactions that may not have been envisioned otherwise.

8.1.1 Applications to Reinforcement Learning, Decision Making and Cognitive Control

The above "recipe" for linking levels of computation is relatively abstract. The rest of this chapter focuses on concrete examples from my lab. We study the neurocomputational mechanisms of cortico-basal ganglia functions—including action selection, reinforcement learning and cognitive control. For theory development, we leverage a combination of two distinct levels of computation. First, our lab and others have simulated interactions within and between corticostriatal circuits via dynamical neural systems models which specify the roles of particular neural mechanisms [27, 28, 31, 35, 44, 59, 74, 76]. These models are motivated by anatomical, physiological, and functional constraints. The mechanisms included in the detailed neural models were originally based on data from animal models, but integrated together into a systems-level functional model (i.e., one that has objectives and links to behavior). As such they have reciprocally inspired rodent researchers to test, and validate, key model predictions using genetic engineering methods in rodents by manipulating activity in separable corticostriatal pathways [45, 51, 67]. Second, we adopt and refine higher level mathematical models to analyze the functional properties of the

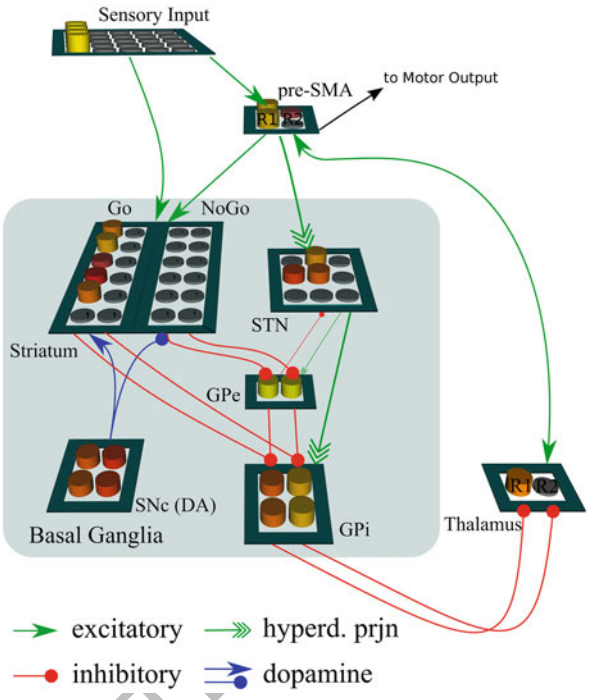

Fig. 8.1 Neural network model of a single cortico-basal ganglia circuit [2]. Sensory and motor (pre-SMA) cortices project to the basal ganglia. Two opposing "Go" and "NoGo" (direct and indirect) pathways regulate action facilitation and suppression based on reward evidence for and against each decision option. Dopamine (DA) modulates activity levels and plasticity in these populations, influencing both choice and learning. The 'hyperdirect' pathway from cortex to STN acts to provide a temporary Global NoGo signal inhibiting the selection of all actions, particularly under conditions of decision conflict (co-activation of competing pre-SMA units). *GPi/e* Globus Pallidus internal/ external segment

neurocognitive systems, affording a principled computational interpretation and allowing for tractable quantitative fits to brain-behavior relationships [5, 13, 22, 23, 59]. Examples of the two levels of description are shown in Fig. 8.1 (neural systems) and Figs. 8.2, 8.3 and 8.4 (abstractions).

For empirical support, we design computerized tasks sensitive to the hypothesized neural computations that probe *reinforcement learning, cognitive control, and reward-based decision making under uncertainty*. We provide quantitative estimates of individual performance parameters using mathematical models, yielding objective assessments of the degree to which subjects rely on specific computations when learning and making decisions [13, 14, 29, 31, 34, 39, 62]. We assess how these parameters vary with markers of neural activity (EEG, fMRI), and how they are altered as a function of illness, brain stimulation, pharmacology, and genetics [13, 14, 34, 42, 54, 58].

This approach has contributed to a coherent depiction of frontostriatal function and testable predictions. As one example, we have identified mechanisms underlying two distinct forms of impulsivity. The first stems from a deficit in learning from "negative reward prediction errors" (when decision outcomes are worse than expected, dependent on dips in dopamine and their resultant effects on activity and plasticity in a subpopulation of striatal neurons expressing D2 dopamine receptors). Deficiencies in the mechanism lead to a failure to properly consider negative outcomes of prospective decisions, and hence lead to a bias to focus primarily on gains.

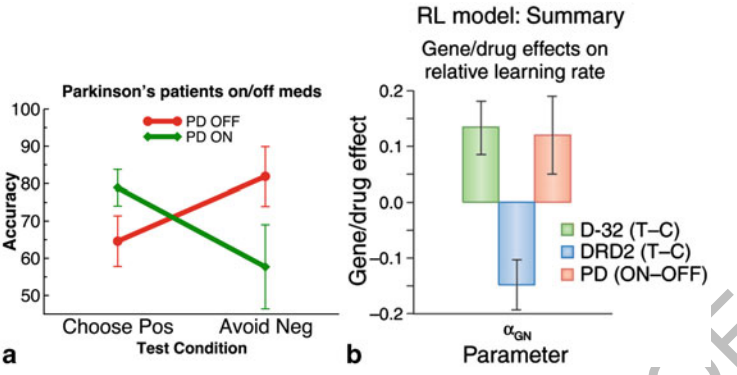

Fig. 8.2 Effects of **a** dopamine medication manipulations in Parkinson's disease [36] and **b** genotypes related to striatal D1 and D2 pathways in healthy participants on choice accuracy [23, 39]. "Choose Pos" assesses the ability to choose the probabilistically most rewarded (positive) action based on previous Go learning; "Avoid Neg" assesses the ability to avoid the probabilistically most punished (negative) action based on previous NoGo learning. **c** Quantitative fits with a reinforcement learning (RL) model capture these choice dissociations by assigning asymmetric Go vs. NoGo learning rates that vary as a function of PD, medications, and genotype [19, 32]. Unlike studies linking candidate genes to complex disease phenotypes, where findings often fail to replicate [56], this linking of neurogenetic markers of corticostriatal function to specific computational processes has been replicated across multiple experiments

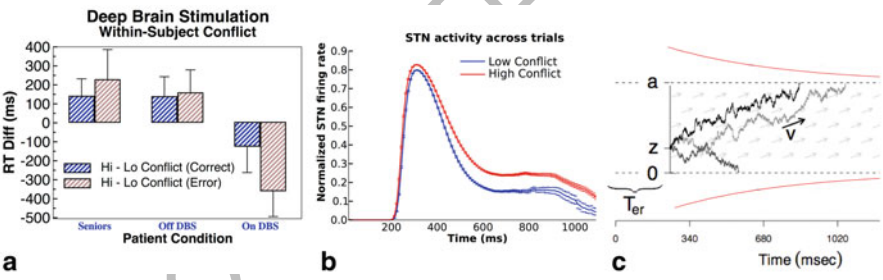

Fig. 8.3 **a** Decision conflict-induced response time slowing in PD patients and controls ("Seniors"). Hi Conflict: alternative actions have similar reward probabilities (high entropy). Lo Conflict: alternative actions have qualitatively different reward probabilities. DBS reverses conflict-induced slowing, leading to impulsive choice (large effect for suboptimal "error" choices) **b** STN firing rate in the neural model surges during action selection, to a greater extent during high conflict trials, delaying responding (not shown). **c** Two-choice decision making is captured by the drift diffusion model. Evidence accumulates for one option over the other from a starting point "z" at a particular average rate "v"; choices are made when this evidence crosses the decision threshold ("a"). Noisy accumulation leads to variability in response times (example RT distributions shown). Dynamics of STN function are captured in this framework by an increased decision threshold when conflict is detected (*red curve*), followed by a collapse to a static asymptotic value, similar to STN activity in (b); see [59]. Quantitative fits of the BG model with the DDM show that STN strength parametrically modulates decision threshold, as supported by experimental manipulations of STN and model-based fMRI studies showing STN activity varying with decision threshold estimated with the DDM [42]

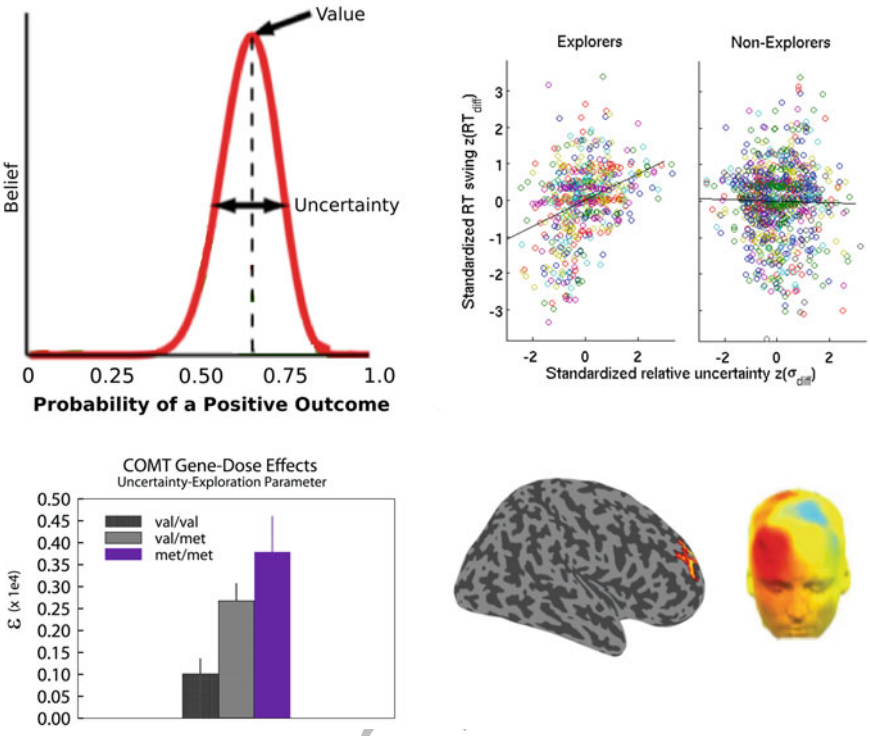

Fig. 8.4 The exploration-exploitation tradeoff. *Top*: probability density functions representing the degree of belief in values of two alternative actions. One action has a higher value estimate, but with uncertainty, quantified by Bayesian updating of belief distributions as a function of experience. Exploration is predicted to occur when alternative actions have relatively higher uncertainty. Approximately half of participants ("Explorers") employ this exploration strategy, with choice adjustments proportional to relative uncertainty. *Bottom*: The degree of uncertainty-driven exploration estimated by model parameter ε varies as a function of a genetic variant in COMT, associated with prefrontal dopamine [41]. fMRI data show that activity in RLPFC varies parametrically with relative uncertainty in Explorers [5, 14]. EEG topographical map shows theta-band activity correlates with relative uncertainty, maximal over rostrolateral electrodes

The second form of impulsivity stems from a failure to adaptively pause the decision process given conflicting evidence about the values of alternative actions (dependent on communication between frontal cortex and the subthalamic nucleus (STN), which temporarily provides a soft brake on motor output). Deficiencies in this mechanism lead to rash choices, i.e., a failure to consider the value of all the decision options but instead to quickly accept or reject an option based on its value alone. Notably these two forms of impulsivity are differentially impacted by distinct forms of treatment in Parkinson's disease –medications that produce elevations in dopamine and deep brain stimulation of the STN—supporting the posited mechanisms by which these treatments alter neural circuitry [13, 14, 39, 80]. Further insights come from linking these precise mechanisms to their high level computations. To do so, we next present

a more detailed overview of the neural model, followed by its linking to abstract computations.

8.2 Cortico-Striatal Interactions During Choice and Learning

The basal ganglia (BG) are a collection of subcortical structures that are anatomically, neurochemically, and functionally linked [1, 29, 44, 55]. Through a network of interconnected loops with the frontal cortex, they modulate motor, cognitive, and affective functions. The defining characteristic of this interaction is a "gating function": the frontal cortex first generates candidate options based on their prior history of execution in the sensory context and the BG facilitates selection [44] of one of these candidates given their relative learned reward values [27, 29].

Choice Two main projection pathways from the striatum go through different striatal nuclei on the way to thalamus and up to cortex. Activity in the *direct* "Go" pathway provides evidence in favor of facilitation of the candidate cortical action, by disinhibiting thalamocortical activity for the action with highest reward value. Conversely, activity in the *indirect* "*NoGo*" pathway indicates that the action is maladaptive and hence should not be gated. Thus for any given choice, direct pathway neurons convey the positive evidence in favor of that action based on learned reward history, whereas indirect pathway neurons signal the negative evidence (likelihood of leading to a negative outcome). The action most likely to be gated is a function of the relative difference in Go/NoGo activity for each action [27]. Dopamine influences the cost/benefit tradeoff by modulating the balance of activity in these pathways via differential effects on D1 and D2 receptors, thereby modulating choice incentive (whether choices are determined by positive or negative potential outcomes). Similar principles apply to the selection of cognitive actions (notably, working memory gating) in BG-prefrontal cortical circuits [27, 29, 35, 42]. Finally, conflict between competing choices, represented in mediofrontal/premotor cortices, activates the subthalamic nucleus (STN) via the *hyperdirect* pathway. In turn, STN activity delays action selection, making it more difficult for striatal Go signals to facilitate a choice and buying more time to settle on the optimal choice [28]; in abstract formulations, this is equivalent to a temporary elevation in the *decision threshold* [59] (see below; Fig. 8.3).

Learning The models simulate phasic changes in dopamine levels that occur during positive and negative *reward prediction errors* (difference between expected and obtained reward), and their effects on plasticity in the two striatal pathways. Phasic bursts of dopamine cell firing during positive prediction errors act as teaching signals that drive Go learning of rewarding behaviors via D1 receptor stimulation [27, 36, 64]. Conversely, negative prediction errors lead to pauses in dopamine firing [64], supporting NoGo learning to avoid unrewarding choices via D2 receptor disinhibition. An imbalance in learning or choice in these pathways can lead to a host of aberrant neurological and psychiatric symptoms [54].

8.3 Higher Level Descriptions

Thus far we have considered basic mechanisms in dynamical models of corticostriatal circuits and their resultant effects on behavior. However, these models are complex (cf. Fig. 8.1—this is the 'base' model and other variants build from there to include multiple circuits and their interactions). Moreover, because they simulate internal neural dynamics consistent with electrophysiological data, they require many more parameters than is necessary to relate to behavior alone. Higher level computational descriptions allow us to abstract away from the detailed implementation. In particular, the tendency to incrementally learn from reward prediction errors and to select among multiple candidate options has been fruitfully modeled using reinforcement learning (RL) models inherited from the computer science literature. These models summarize valuation of alternative actions, reflecting the contributions of the striatal units in the neural models, in terms of simple "Q values" which are incremented and decremented as a function of reward prediction errors. A simple choice function is used to compare Q values across all options and to stochastically select that with the highest predicted value: this function summarizes the effective computations of the gating circuitry that facilitates cortical actions (where noise in both cortex and striatum results in stochastic choice function). An asymmetry in learning from positive or negative prediction errors (reflecting high or low dopamine levels and their effects on activity/plasticity) can be captured by using separate learning rates (Frank et al. 2007) [23]. However, a better depiction of the neural model allows for not only differential modulation of learning, but also differential modulation of choice incentive during action selection [19]. This model uses separate Q values, QG and QN, to represent the Go and NoGo pathway respectively, each with their own learning rate, and where the 'activity', i.e. the current QG or QN value, further influences learning. The choice probability is then a function of the relative difference in QG and QN values for each decision option, with a gain factor that can differentially weigh influences of QG vs QN. This gain factor can be varied to simulate dopamine effects on choice incentive by boosting the extent to which decisions are made based on learned QG or QN values, even after learning has taken place.

Because these models are simple and minimal, they can be used to quantitatively fit behavioral data, and to determine whether the best fitting parameters vary as a function of biological manipulations. But because there is a clear mapping from these models to the neural versions, there are strong *a priori* reasons to manipulate particular neural mechanisms and to test whether the resulting estimates of computational parameters are altered as predicted or not.

Indeed, evidence validating these multi-level model mechanisms has mounted over the last decade across species. Monkey recordings combined with Q learning model fits indicate that separate populations of striatal cells code for positive and negative Q values associated with action facilitation and suppression [24, 52, 62, 72]. In mice, targeted manipulations confirm selective roles of direct and indirect pathways in the facilitation and suppression of behavior [50], which are necessary and sufficient to induce reward and punishment learning, respectively [45, 51]. Phasic

stimulation or inhibition of dopamine neurons induces reward/approach and aversive/avoidance learning, respectively [68, 69]. Synaptic plasticity studies reveal dual mechanisms for potentiation and depression in the two pathways as a function of D1 and D2 receptors [65], as in the models. In humans, striatal dopamine manipulation influences the degree to which individuals learn more from positive or negative outcomes (Fig. 8.2), with DA elevations enhancing reward learning but impairing punishment learning, and vice-versa for DA depletion [6, 34, 36, 58, 70] (Frank et al. 2007). Quantitative fits using RL models reveal that these can be accounted for by differential effects of dopamine manipulations on learning rates from positive and negative prediction errors. Moreover, in the absence of any acute manipulation, individual differences in these fit learning rate parameters are associated with genetic polymorphisms that differentially impact the efficacy of striatal D1 and D2 pathways [23, 32, 33, 41, 42] (Fig. 8.2).

One of the advantages of high level models, besides being simpler and more naturally used for quantitative behavioral fits, is that they can also include relevant processes that are out of scope in the neural versions. For example, when humans perform a "reinforcement learning task", they are not only incrementally learning probabilistic stimulus-action-outcome associations and choosing between them, but they also engage in other cognitive strategies involving hypothesis testing and working memory. Fitting their behavior with a RL model alone—no matter how well this model summarizes the corticostriatal learning process and its contribution to behavior—is then misleading, because it will capture variance that is really due to working memory capacity by absorbing this into the learning rate parameters of the RL process. Collins and Frank [17] showed clear evidence of such effects by manipulating the number of stimuli in the set to be learned (and hence working memory load). They found that when using RL models alone and without factoring in working memory, one needed to include a separate learning rate for each set size to capture the data, and that a gene related to prefrontal but not striatal function was predictive of this learning rate. However, when an augmented model which included a capacity-limited working memory process was used, the overall fits to the data were improved, and the RL process could be captured by a single learning rate that applies across all set sizes. Further, this learning rate in this best fit model varied with striatal genetic function, whereas the prefrontal gene was now related to working memory capacity.

On the other hand, algorithmic RL models that only predict choice probability miss out on the dynamics of choice, reflected in RT distributions, which emerge naturally from the neural model because it is a *process model*. First, firing rate noise throughout the network produces variance in the

speed with which an action is gated. Second, the action value of the candidate option impacts not only the likelihood of selecting that option relative to its competitors, but also the speed with which this option is selected. Finally, as mentioned above, when multiple candidate options have similar frequencies of execution based on their choice history—that is, when there is conflict or choice entropy—this elicits hyperdirect pathway activity from mediofrontal cortex to the STN, which provides a

temporary brake on the striatal gating process, thereby slowing down response time and increasing the likelihood in settling on the optimal response [28].

High level descriptions of process models have been extensively used to simulate dynamics of simple decision making in cognitive psychology for over 3 decades. In particular, the drift diffusion model (DDM) belongs to a class of sequential sampling models in which noisy evidence is accumulated in favor of one of two options, and a choice is executed once this evidence cross a critical *decision threshold*. The slope at which evidence accumulates is called the drift rate and reflects the ease of the decision. These models capture not only choice proportions and mean RT, but the entire shape of the RT distribution for correct and erroneous responses.

Notably, when fitting the behavioral outputs of the neural model with the DDM, we found that parametric manipulations of both corticostriatal and STN output projection strengths were related to estimated decision threshold, with corticostriatal strength decreasing threshold (see [25]) and STN strength increasing threshold [59, 75].

Studies with Parkinson's patients on and off STN deep brain stimulation provide an opportunity to test the impact of interference of the STN pathway, which can also lead to clinical impulsivity. Indeed, this procedure provides a selective disruption of conflict-induced slowing, without impacting learning [20, 40] (Fig. 8.3). We have extended this finding in three critical ways. First, EEG revealed that in healthy participants and patients off DBS, the amount of medial prefrontal (mPFC) theta-band activity during high conflict trials was predictive on a trial-to-trial basis of the amount of conflict-induced RT slowing. STN-DBS reversed this relationship, presumably by interfering with hyperdirect pathway function, without altering mPFC theta itself. Second, we developed a toolbox for hierarchical Bayesian parameter estimation allowing us to estimate the impact of trial-to-trial variations in neural activities on decision parameters [78]. We found that mPFC theta was predictive of decision threshold adjustments (and not other decision parameters), and, moreover, that DBS reversed this mPFC-threshold relationship [13, 14]. Third, electrophysiological recordings within STN revealed decision conflict-related activity in a similar time and frequency range as mPFC in both humans and monkeys [4, 13, 14, 42, 46, 47, 80]. These findings thus provide support for a computational account of hyperdirect pathway function, and a potential explanation for the observed impulsivity that can sometimes result from DBS.

Thus far we have considered the ability of existing abstract formulations to summarize the computations of more detailed neural models, providing a link between levels. It is also possible however, that aspects of the neural models, if valid, should alter the way we think about the abstract formulation. In the above example, we claimed that the STN was involved in regulating decision threshold. Consider its internal dynamics however (Fig. 8.3b). STN activity is not static throughout a trial, but rather exhibits an initial increase in activity, which then subsides with time during the action selection process. Moreover the initial STN surge is larger and more prolonged when there is higher decision conflict. This model dynamic is supported by electrophysiological evidence in both monkeys and humans [47, 80], and implies that STN effects on preventing BG gating should be transient and decrease with time, implying a collapsing rather than fixed decision threshold. Functionally this

collapsing threshold ensures that a decision is eventually made, preventing decision paralysis (this collapsing threshold is optimal when there are response deadlines; [43]). Indeed, quantitative fits using the DDM to capture RT distributions of the BG model showed that a collapsing threshold provided a good account of the model's behavior, notably, with the temporal dynamics of the best fitting exponentially collapsing threshold matching reasonably well to the dynamics of STN activity—despite the fact that the DDM fits had no access to this activity but only to RT distributions [59]. This study also found that when fitting human behavioral data in the same reward conflict decision-making task, fits were improved when assuming a higher and collapsing threshold in conflict trials, compared to the fixed threshold model.

This last result supports the assertion that neural mechanism constraints can be included to refine higher level descriptions. However, we must also admit that we do not have well constrained neural mechanistic models for all cognitive processes. The next example I turn to is the exploration-exploitation tradeoff in reinforcement learning, a process studied in machine learning for many years but only recently considered in the cognitive neurosciences.

8.4 Beyond Basic Mechanisms: Uncertainty Driven Exploration and Hierarchical Learning

Often individuals need to explore alternative courses of action to maximize potential gains. *But how does one know when to explore rather than exploit learned values?* Basic RL models usually assume a degree of random exploration, but a more efficient strategy is to keep track of the uncertainty about value estimates, and to guide exploration toward the action with higher uncertainty [21]. We have reported evidence for just such a mechanism, whereby trial-by-trial behavioral adjustments are quantitatively related to a Bayesian model estimate of relative outcome uncertainty. In this case, there is no existing neural model for how this relative uncertainty measure is encoded or updated as a function of reward experiences. Nevertheless, individual differences in the employment of this uncertainty-driven exploration strategy are predicted by genetic variations in the COMT (Catechol-O-methyltransferase) gene, which is related to prefrontal cortical dopamine function [41]. Further, a recent model-based fMRI study [5] revealed that the rostrolateral prefrontal cortex (RLPFC) parametrically tracks the relative uncertainty between outcome values, preferentially so in "Explorers" (defined based on behavioral fits alone). In EEG, relative uncertainty is reflected by variations in theta power over RLPFC (in contrast to the mPFC indices of conflict noted above), again preferentially in Explorers [13]. These converging data across modeling, behavior, EEG, genetics and fMRI indicate a potential prefrontal strategy for exploring and over-riding reward-based action selection in the BG. Notably, patients with schizophrenia, specifically those with anhedonia, exhibit profound reductions in uncertainty-driven exploration [66]. Thus this measure has potential relevance for understanding motivational alterations in clinical populations,

and motivates the development of mechanistic models of how relative uncertainty estimates are computed and updated in populations of prefrontal neurons.

As the field matures, it becomes less clear which level of modeling motivated the other—and this is a good thing, as mutual constraints become available. Collins and Frank [19] confronted the situation in which a learner has to decide whether, when entering a new context, the rules dictating links between states, actions and outcomes ("task-sets") should be re-used from those experienced in previous contexts, or whether instead a new task-set should be created and learned.

They developed a high level "context-task-set" (C-TS) computational model based on non-parametric Bayesian methods (Dirichlet process mixtures), describing how the learner can cluster contexts around task-set rules, generalizable to novel situations. This model was motivated by analogous clustering models in category learning (e.g., [1, 63]), but applied to hierarchical cognitive control, and as such was similarly motivated by the hierarchical structure of prefrontal cortical basal ganglia networks and modeling implementations thereof [10, 30, 49, 61]. They also constructed a refined hierarchical PFC-BG network which confronted the same tasks, and showed that its functionality is well mimicked by the C-TS model. Quantitative model fitting linking these levels showed that particular neural mechanisms were associated with specific C-TS model parameters. For example, the prior tendency to re-use vs. create new structure in C-TS, captured by Dirichlet alpha parameter, was directly related to the sparseness of the connectivity matrix from contextual input to PFC (Fig. 8.5). Thus in this case, there existed well established and validated models of interactions between PFC and BG during learning, working memory, and action selection (including some hierarchical implementations), but the computations afforded by the novel C-TS model further inspired refinement and elaboration of the network. In turn, this exercise reciprocally allowed us to derive more specific predictions about mechanisms leading to differential response times and error patterns (which were confirmed behaviorally), and to marry the reinforcement learning models described previously with the cognitive control mechanisms involving decision threshold regulation.

One novel finding from this modeling work was that in such environments, the STN mechanism, previously linked only to decision making and impulsivity, plays a key role in learning. In particular, the simulations showed that early in the trial, when there is uncertainty about the identity of the PFC task-set, this conflict between alternative PFC states activated the STN, preventing the motor loop from responding. This process ensures that the PFC state is resolved prior to motor action selection, and as such, when the outcome arrives, stimulus-action learning is conditionalized by the selected PFC state. As the STN contribution is reduced, there is increasing interference in learning across task-sets, hence learning is less efficient. This novel theory specifying the role of the STN in conditionalizing learning by PFC state needs to be tested empirically (e.g. with DBS or fMRI), but each component is grounded by prior empirical and theoretical work, yet it would not likely have emerged without this multi-level modeling endeavor. In related work, Frank and Badre [30] considered hierarchical learning tasks with multidimensional stimuli with two levels of modeling. A Bayesian mixture of experts model summarized how participants may learn hierarchical structure of the type, "if the color is red, then the response is determined

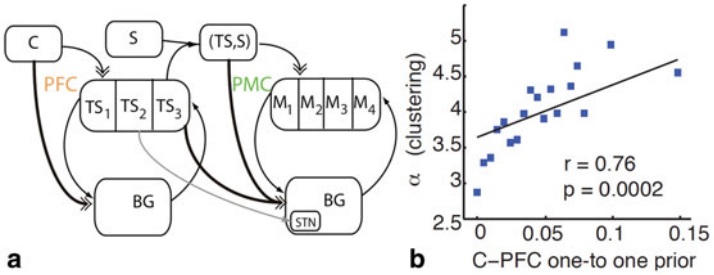

Fig. 8.5 *Left*: Schematic of hierarchical corticostriatal network model for creating task-sets (TS) which are gated into prefrontal cortex depending on the context C. The lower motor loop selects motor actions M depending on the selected TS in PFC and the current sensory state S. The same TS can be reused across contexts, supporting clustering and generalization of behaviors, or if needed, a new TS can be gated, preventing interference in learned state-action mappings between different contexts/TS. *Right*: Parametric manipulation of the sparseness of the connectivity matrix from Context to PFC (enforcing a prior tendency to encode distinct C's as distinct PFC TS) is well fit by an increased α Dirichlet process clustering parameter in the C-TS model which creates and re-uses TS according to non-parametric Bayesian methods. (Adapted from Collins and Frank [18])

by the shape, whereas if the color is blue, the response is determined by the orientation." Quantitative modeling showed that estimated attention to the hierarchical expert was linked to speeded learning in hierarchical conditions, and when fit to a PFC-BG network, was related to a measure of gating policy abstraction, learned via RL, in the hierarchical connections from PFC to striatum. Badre and Frank (2012) then used model-based fMRI to show that in participants, estimated attention to hierarchical structure was linked to PFC-BG activity within a particular rostrocaudal level of the network consistent with the "second-order" rule level of the task.

8.5 Concluding Comments

The examples described above demonstrate mutual, reciprocal constraints between models of neural circuitry and physiology to models of computational function. This exercise leads to multiple testable predictions using model-based cognitive neuroscience methods. Ultimately, models are judged based on their predictive power, and as such, they can inspire informative experiments valuable even to those who question the validity or assumptions of either of the levels of modeling employed.

8.6 Exercises

1. Give examples of implementational neural models and higher level algorithmic models in any domain. What sorts of data do these models attempt to capture?
2. Think of some examples in which an abstract model exists but would benefit from a mechanistic elaboration for making cognitive neuroscience predictions.

3. Can you think of potential advantages of combining the models? How about some pitfalls?
4. Describe how dopamine may contribute both to learning and to choice incentive (the degree to which decisions are made based on positive vs negative consequences).
5. Reinforcement learning models and sequential sampling models of decision making have been largely separate literatures in mathematical psychology yet each of these classes of models have been fit to the basal ganglia neural model described in this chapter. Read Bogacz and Larsen [8] for a complementary approach to linking these formulations within an algorithmic framework.
6. Conversely, read Wong and Wang [79] for a complementary example of a single neural model capturing dynamics of decision making and working memory.

8.7 Solutions

1. Daniel Durstewitz has several detailed neural models of prefrontal dopamine mechanisms in working memory. These are complemented by algorithmic models of working memory updating (e.g. see [16] for a discussion of both levels). In this case, the Durstewitz models capture attractor dynamics and effects of D1 receptors on sodium and potassium currents, etc, whereas the algorithmic models simulate performance in working memory tasks as a function of reward prediction errors.
2. Uncertainty driven exploration (see Sect. 5), as one example
3. Advantages discussed in this chapter. Pitfalls: perhaps assumptions of either level of modeling are flawed, and one might actually detract from the other.
4. See Sects. 3 and 4.
5. –
6. –

8.8 Further Reading

1. Collins and Frank [18] present two levels of modeling describing the interactions between cognitive control and learning needed to construct task-set rules generalizable to novel situations. This endeavor reaps the benefits of both RL models and the temporal dynamics of decision making, and how each affects the other. It also shows theoretically how a non-parametric Bayesian approach to task-set clustering can be implemented in hierarchical PFC-BG circuitry.
2. Wang [71] reviews neural models of decision making and their relation to normative theory.
3. Brittain et al. [12] present evidence for STN involvement in deferred choice under response conflict in a non-reward based task, complementing findings described in this chapter.

References

1. Alexander GE, Crutcher MD (1990) Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci 13(7):266–271
2. Anderson JR (1991) The adaptive nature of human categorization. Psychol Rev 98(3):409–429
3. Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack R (2007) Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. J Neurosci 27(14):3743–3752
4. Badre D, Frank MJ (2012) Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 2: Evidence from fMRI. Cerebral Cortex 22:527–536
5. Badre D, Doll BB, Long NM, Frank MJ (2012) Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron 73:595–607
6. Bódi N, Kéri S, Nagy H, Moustafa A, Myers CE, Daw N, Dibó G, Takáts A, Bereczki D, Gluck MA (2009). Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain 132:2385–2395
7. Bogacz R, Gurney K (2007) The basal ganglia and cortex implement optimal decision making between alternative actions. Neural Comput 19(2):442–477
8. Bogacz R, Larsen T (2011) Integration of reinforcement learning and optimal decision-making theories of the basal ganglia. Neural Comput 23(4):817–851
9. Bogacz R, Wagenmaker EJ, Forstmann BU, Nieuwenhuis S (2010) The neural basis of the speed-accuracy tradeoff. Trends Neurosci 33(1):10–16
10. Botvinick MM, Niv Y, Barto AC (2009) Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. Cognition 113(3):262–280
11. Braver TS, Cohen JD (2000) On the control of control: the role of dopamine in regulating pre- frontal function and working memory. In: Monsell S, Driver J (eds) Control of cognitive processes: attention and performance XVIII. MIT Press, Cambridge, pp 713–737
12. Brittain JS, Watkins KE, Joundi RA, Ray NJ, Holland P, Green AL, Aziz TJ, Jenkinson N (2012) A role for the subthalamic nucleus in response inhibition during conflict. J Neurosci 32(39):13396–13401
13. Cavanagh JF, Frank MJ, Klein TJ, Allen JJB (2010) Frontal theta links prediction error to behavioral adaptation in reinforcement learning. Neuroimage 49(4):3198–3209
14. Cavanagh JF, Figueroa CM, Cohen MX, Frank MJ (2011a) Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. Cereb Cortex 22(11):2575–2586
15. Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ (2011b) Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat Neurosci 14(11):1462–1467
16. Cohen JD, Braver TS, Brown JW (2002) Computational perspectives on dopamine function in prefrontal cortex. Curr Opin Neurobiol 12(2):223–229
17. Collins AGE, Frank MJ (2012) How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. Eur J Neurosci 35(7):1024–1035
18. Collins AGE, Frank MJ (2013) Cognitive control over learning: creating, clustering and generalizing task-set structure. Psychol Rev 120(1):190–229
19. Collins AGE, Frank MJ (2014) Opponent Actor Learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev 121:337–366
20. Coulthard EJ, Bogacz R, Javed S, Mooney LK, Murphy G, Keeley S, Whone AL (2012). Distinct roles of dopamine and subthalamic nucleus in learning and probabilistic decision making. Brain 135:3721–3734
21. Dayan P, Sejnowksi T (1996) Exploration bonuses and dual control. Mach Learn 25:5–22
22. Doll BB, Jacobs WJ, Sanfey AG, Frank MJ (2009) Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. Brain Res 1299:74–94
23. Doll BB, Hutchison KE, Frank MJ (2011) Dopaminergic genes predict individual differences in susceptibility to confirmation bias. J Neurosci 31(16):6188–6198

24. Ford KA, Everling S (2009) Neural activity in primate caudate nucleus associated with pro- and antisaccades. J Neurophysiol 102(4):2334–2341
25. Forstmann BU, Dutilh G, Brown S, Neumann J, von Cramon DY, Ridderinkhof KR, Wagenmakers EJ (2008a) Striatum and pre-SMA facilitate decision-making under time pressure. Proc Natl Acad Sci USA 105(45):17538–17542
26. Forstmann BU, Jahfari S, Scholte HS, Wolfensteller U, van den Wildenberg WP, Ridderinkhof KR (2008b) Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. J Neurosci 28(39):9790–9796
27. Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J Cogn Neurosci 17(1):51–72
28. Frank MJ (2006) Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw 19(8):1120–1136
29. Frank MJ (2011) Computational models of motivated action selection in corticostriatal circuits. Curr Opin Neurobiol 2:381–386
30. Frank MJ, Badre D (2012) Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb Cortex 22(3):509–526
31. Frank MJ, Claus ED (2006) Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol Rev 113(2):300–326
32. Frank MJ, Fossella JA (2011) Neurogenetics and pharmacology of learning, motivation, and cognition. Neuropsychopharmacology 36:133–152
33. Frank MJ, Hutchison K (2009) Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. Neuroscience 164(1):131–140
34. Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behav Neurosci 120(3):497–517
35. Frank MJ, Loughry B, O'Reilly RC (2001) Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cogn Affect Behav Neurosci 1(2):137–160
36. Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306(5703):1940–1943
37. Frank MJ, Santamaria A, O'Reilly R, Willcutt E (2007a) Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. Neuropsychopharmacology 32(7):1583–1599
38. Frank MJ, D'Lauro C, Curran T (2007b) Cross-task individual differences in error processing: neural, electrophysiological, and genetic components. Cogn Affect Behav Neurosci 7(4):297–308
39. Frank MJ, Moustafa AA, Haughey H, Curran T, Hutchison K (2007c) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci 104(41):16311–16316
40. Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007d) Hold your horses: impulsivity, deep brain stimulation and medication in Parkinsonism. Science 318:1309–1312
41. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat Neurosci 12(8):1062–1068
42. Frank MJ, Gagne C, Nyhus E, Masters S, Wiecki TV, Cavanagh JF, Badre D (2015) fMRI and EEG Predictors of dynamic decision parameters during human reinforcement learning. J Neurosci 35
43. Frazier P, Yu AJ (2008) Sequential hypothesis testing under stochastic deadlines. Adv Neural Inf Process Syst 20:465–472. (MIT Press, Cambridge)
44. Gurney K, Prescott TJ, Redgrave P (2001) A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. Biol Cybern 84(6):411–423
45. Hikida T, Kimura K, Wada N, Funabiki K, Nakanishi S (2010) Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. Neuron 66(6):896–907

46. Isoda M, Hikosaka O (2007) Switching from automatic to controlled action by monkey medial frontal cortex. Nat Neurosci 10(2):240–248
47. Isoda M, Hikosaka O (2008) Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. J Neurosci 28(28):7209–7218
48. Jahfari S, Verbruggen F, Frank MJ, Waldorp LJ, Colzato L, Ridderinkhof KR, Forstmann BU (2012) How preparation changes the need for top-down control of the basal ganglia when inhibiting premature actions. J Neurosci 32(32):10870–10878
49. Koechlin E, Summerfield C (2007) An information theoretical approach to the prefrontal executive function. Trends Cogn Sci 11(6):229–235
50. Kravitz AV, Freeze BS, Parker PRL, Kay K, Thwin MT, Deisseroth K, Kreitzer AC (2010) Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature 466(7306):622–626
51. Kravitz AV, Tye LD, Kreitzer AC (2012) Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nat Neurosci 15:816–818
52. Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching behavior. Neuron 58(3):451–463
53. Lo CC, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat Neurosci 9(7):956–963
54. Maia TV, Frank MJ (2011) From reinforcement learning models to psychiatric and neurological disorders. Nat Neurosci 2:154–162
55. Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol 50(4):381–425
56. Munafò MR, Stothart G, Flint J (2009) Bias in genetic association studies and impact factor. Mol Psychiatry 14:119–120
57. O'Reilly RC, McClelland JL (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. Hippocampus 4(6):661–682
58. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442(7106):1042–1045
59. Ratcliff R, Frank MJ (2012) Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. Neural Comput 24:1186–1229
60. Ratcliff R, McKoon G (2008) The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput 20(4):873–922
61. Reynolds JR, O'Reilly RC (2009) Developing PFC representations using reinforcement learning. Cognition 113(3):281–292
62. Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310(5752):1337–1340
63. Sanborn AN, Griffiths TL, Navarro DJ (2010) Rational approximations to rational models: alternative algorithms for category learning. Psychol Rev 117(4):1144–1167
64. Schultz W (2002) Getting formal with dopamine and reward. Neuron 36(2):241–263
65. Shen W, Flajolet M, Greengard P, Surmeier DJ (2008) Dichotomous dopaminergic control of striatal synaptic plasticity. Science 321(5890):848–851
66. Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, Gold JM (2011) Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. Biol Psychiatry 69:424–431
67. Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012) Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nat Neurosci 15:1281–1289
68. Tan KR, Yvon C, Turiault M, Mirabekov JJ, Doehner J, Labouèbe G, Deisseroth K, Tye KM, Lüscher C (2012) GABA neurons of the VTA drive conditioned place aversion. Neuron 73:1173–1183
69. Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science 324:1080–1084
70. Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, Hallett M (2010) Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. Neuron 65(1):135–142

71. Wang XJ (2012) Neural dynamics and circuit mechanisms of decision-making. Curr Opin Neurobiol 22:1039–1046
72. Watanabe M, Munoz DP (2009) Neural correlates of conflict resolution between automatic and volitional actions by basal ganglia. Eur J Neurosci 30(11):2165–2176
73. Watkins CJCH, Dayan P (1992) Q-Learning. Mach Learn 8:279–292
74. Wiecki TV, Frank MJ (2010) Neurocomputational models of motor and cognitive deficits in Parkinson's disease. Prog Brain Res 183:275–297
75. Wiecki TV, Frank MJ (in press). A computational model of inhibitory control in frontal cortex and basal ganglia. Psychological Review.
76. Wiecki TV, Riedinger K, Meyerhofer A, Schmidt W, Frank MJ (2009) A neurocomputational account of catalepsy sensitization induced by D2 receptor blockade in rats: context dependency, extinction, and renewal. Psychopharmacology (Berl) 204:265–277
77. Wiecki TV, Sofer I, Frank MJ (2012). Hierarchical Bayesian parameter estimation of Drift Diffusion Models (Version 0.4RC1) [software]. http://ski.clps.brown.edu/hddm_docs/.
78. Wiecki TV, Sofer I, Frank MJ (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. Fron Neuroinformatics 7:1–10
79. Wong KF, Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. J Neurosci 26(4):1314–1328
80. Zaghloul K, Weidemann CT, Lega BC, Jaggi JL, Baltuch GH, Kahana MJ (2012) Neuronal activity in the human subthalamic nucleus encodes decision conflict during action selection. J Neurosci 32(7):2453–2460