6

Computational Cognitive Neuroscience Approaches to Deconstructing Mental Function and Dysfunction

Michael J. Frank

Abstract

Advances in our understanding of brain function and dysfunction require the integration of heterogeneous sources of data across multiple levels of analysis, from biophysics to cognition and back. This chapter reviews the utility of computational neuroscience approaches across these levels and how they have advanced our understanding of multiple constructs relevant for mental illness, including working memory, reward-based decision making, model-free and model-based reinforcement learning, exploration versus exploitation, Pavlovian contributions to motivated behavior, inhibitory control, and social interactions. The computational framework formalizes these processes, providing quantitative and falsifiable predictions. It also affords a characterization of mental illnesses not in terms of overall deficit but rather in terms of aberrations in managing fundamental trade-offs inherent within healthy cognitive processing.

Introduction

Understanding how any system, including the brain, can become dysfunctional first requires at least a general understanding of how it is functional. Arguably the main obstacle to progress in psychiatry is its historical inclination to "put the cart before the horse" in its efforts to link illnesses and higher-level symptoms to individual genes or molecular mechanisms, before understanding the relationship between the many intermediate levels of analysis. Cognitive neuroscience takes smaller (but still lofty) leaps of linking isolated cognitive processes to larger-scale mechanisms (often with coarse descriptors) to offer better explanations of neurocognitive processes, but with less immediate application

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

to mental illness. One of the central goals of the computational psychiatry mission is to develop principled mechanistic models that formalize functional objectives within the domains of perception, action, and cognition, and to explore how aberrations in such mechanisms lead to corresponding changes in mental function. As such, in this chapter I discuss only selected computational cognitive neuroscience approaches and domains relevant for mental illness, focusing on high-level concepts rather than details of the formulations. For other treatments of these thematic ideas and related assumptions and practices of computational psychiatry, see Maia and Frank (2011), Montague et al. (2012), Huys et al. (2012), Stephan and Mathys (2014), Huys et al. (2015a), Wiecki et al. (2015), and Wang and Krystal (2014).

Systems and cognitive neuroscience inherently link levels of analysis across a continuum: from biological mechanism to cognitive and behavioral phenomena. One can study, for example, ion channel conductances and receptors, intracellular signaling cascades, synaptic plasticity, excitation-inhibition balance, or probabilistic population codes. Zooming out a level, entire neural systems can be studied, such as the sensory cortices, frontal cortex, hippocampus and basal ganglia, ascending neuromodulatory signals, and interactions among several of these systems. At the cognitive level, a different set of topics and principles are relevant, for example, concerning attention, working memory, decision making, cognitive control, reinforcement learning, and episodic memory. A key role for computational models is their ability to provide a unifying coherent framework that links these levels, specifying computational objectives of a cognitive problem and providing novel interpretation of underlying mechanisms, while also forcing one to be explicit about the assumptions being made.

Indeed, computational approaches encompass a huge range: from those that specify detailed biophysics to those that consider high-level goals of a functional system without regard for implementation. Biophysical models explore, for example, how combinations of ionic currents and their dynamics give rise to higher-level "behaviors," but where behavior here is often defined in terms of the changes in membrane potentials of individual neurons or distinct compartments within neurons, or in terms of synchrony of neural firing across populations of cells. Sometimes these models explore further the impact of particular ionic currents on attractor dynamics thought to be relevant for cognitive function (but usually without simulating realistic cognitive tasks). The field of computational cognitive science, on the other hand, considers how behavioral phenomena might be interpreted—independently of the neural implementation—as optimizing some computational goal, such as 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 approximate (or even directly implement) optimal solutions to computational problems, and how

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

alterations in these mechanisms can lead to predictable changes in behavior. Even here there exist several levels of models: neural network models typically capture some aspects of electrophysiology and dynamics whereas higher-level algorithmic models can summarize the key processes with few free parameters and are more suitable to quantitatively fit behavioral data. By linking the mechanisms at the neural-level model to observable changes in higher-level model parameters (for a review, see Frank 2015), one can derive predictions for how changes in neural activity due to disease, medication, or brain stimulation results in changes in cognitive computations. This linking process also imposes mutual constraint relations between higher- and lower-level descriptions, and allows both levels to be refined and/or reinterpreted by the other.

Moreover, one of the central aims of computational cognitive neuroscience is to identify *computational trade-offs* inherent in cognitive problems and to examine how the brain mitigates these trade-offs at the systems level. As one classic example shows, in memory there is a trade-off between being able to separately store distinct events (e.g., the location of where I parked my car today compared to yesterday) versus being able to accumulate information across events into a coherent representation (e.g., determining the best parking strategy on average). The former process requires distinct neural patterns to encode distinct events, whereas the latter requires a shared population of neurons representing all the times a particular strategy was used to associate with outcomes. Computational cognitive neuroscience approaches have suggested that the brain solves this trade-off by incorporating multiple memory systems in the hippocampus and neocortex (McClelland et al. 1995). Moreover, within the hippocampus, subregions such as the dentate gyrus can enhance pattern separation to minimize interference among similar events, whereas area CA3 supports pattern completion to allow retrieval of memories given partial cues (McClelland et al. 1995). Such models have inspired decades of research and empirical data that have confirmed their key predictions, which may not have been tested in absence of guiding theory, and served to refine further model development.

In what follows I give a brief survey of computational cognitive neuroscience approaches to select problems, highlighting various trade-offs that may be informative for mental illness.

Working Memory and Prefrontal Cortex

The prefrontal cortex (PFC) is long known to be involved in working memory, that is, the ability to hold task-relevant information in mind over short periods of time and use that information to contextualize and guide subsequent actions. Biophysical models have shown that the ability to sustain stable, persistent PFC activation states related to working memory depends on recurrent excitation, intracellular ionic currents and NMDA receptors, modulated by dopaminergic

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

input (Figure 6.1) (for reviews, see Durstewitz and Seamans 2008; Wang and Krystal 2014). These models have made precise predictions regarding the effects of NMDA and dopaminergic manipulations and how they affect attractor states needed for working memory, and have also been instrumental in guiding study on the neural basis of working memory impairments in schizophrenia, associated with NMDA and dopamine (DA) hypofunction in the PFC.

If PFC neurons can maintain information in working memory, a related question is: What is the representational content of this information? Traditional analysis focuses on the coding of individual stimulus dimensions (e.g., color) or basic rule representations. Recent computational models and analysis have instead highlighted that a large proportion of PFC neurons have mixed selectivity across sensory dimensions; that is, their representations are rather high dimensional. Modeling suggests that this "multiplexing" of multiple variables affords a computational advantage by increasing the repertoire of possible input–output mappings that can be read out of attractor states (Rigotti et al. 2013). Support for this interpretation came from an analysis of trials in which monkeys committed cognitive errors (responding according to the wrong rule, thus permitting the analysis of failure modes); in these cases, the dimensionality of PFC neurons collapsed while the simpler low-dimensional representations of individual cues remained intact.

However, computational approaches have also identified a key trade-off within working memory: while it is desirable to robustly maintain task-relevant information in the face of distracting interference (noise in neural firing, task-irrelevant environmental input), it is also important to be able to flexibly update working memory states when incoming information is relevant or when behavioral strategies and plans need to be adjusted. The ability to both robustly maintain existing attractors and rapidly update them are at odds with each other but can be solved by a gating mechanism that dynamically increases the influence of incoming information. Multilevel models have identified DA as a likely candidate for implementing this gating function: when incoming information is relevant, phasic increases in DA can shift the balance in PFC from a "D1 state" optimized for robust maintenance, which is difficult to destabilize, to a "D2 state" optimized for flexibility in terms of allowing for shifting representations (Cohen et al. 2002; Durstewitz and Seamans 2008). Biophysical models show how this pattern emerges due to differential D1 versus D2 receptor sensitivity to different levels of DA, and their differential effects on NMDA and GABA currents, and their resulting network effects on attractor dynamics within populations of cortical pyramidal cells and interneurons.

This example serves to highlight that more DA does not imply better overall function, but rather that a dynamic range of DA signaling is needed to dynamically modulate a functional objective that balances a trade-off: facilitating a switch from maintenance to updating in the service of relevant task demands. Such insights have informed a variety of empirical evidence that DA and PFC states trade off in tasks that demand cognitive stability versus flexibility (Cools

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

From "Computational Psychiatry: New Perspectives on Mental Illness," A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

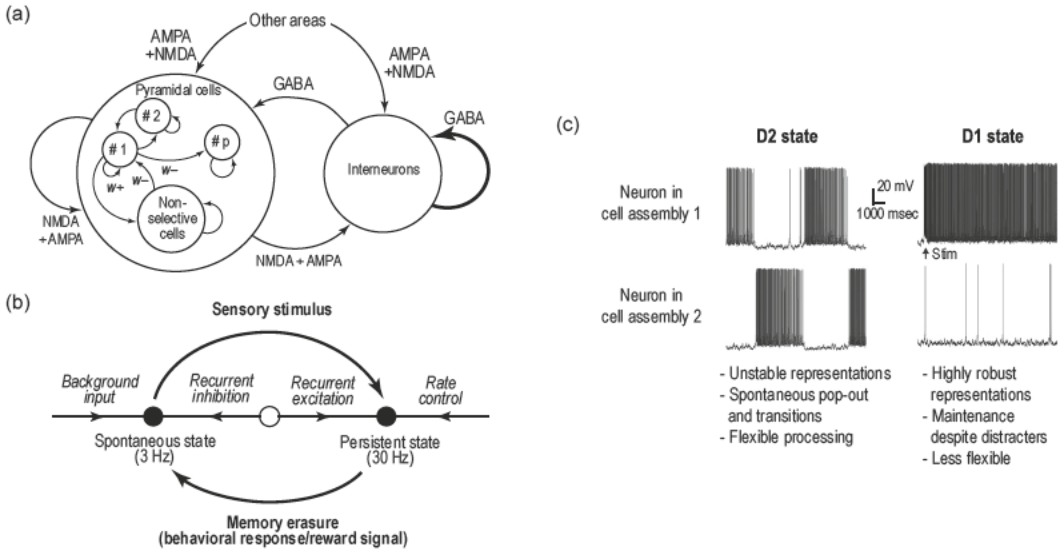

Figure 6.1 Prefrontal cortex and working memory. (a) Biophysical models of prefrontal cortex include multiple populations of pyramidal cells, their interactions with inhibitory interneurons, the differential effects of AMPA and NMDA currents, and recurrent connectivity between these cell populations leading to *attractor dynamics*. Reprinted with permission from Brunel and Wang (2001). (b) Schematic of attractor dynamics for working memory. Background input supports a spontaneous spike rate of prefrontal cells regulated by recurrent inhibition. A potent sensory stimulus provides sufficient excitation to elicit a higher spike rate, which is supported by recurrent excitation leading to a persistent state (*active maintenance).* Other factors can determine when such representations are no longer needed (e.g., task end state reached, behavioral response, reward), driving the population back to spontaneous state. Reprinted with permission from Wang (2001). (c) Biophysical models of dopaminergic influences show how dopamine modulates such prefrontal attractor dynamics. In the D2 state, dopamine levels are high and spiking in model neurons support multiple representations, facilitating updating of working memory. In contrast, the D1 state promotes the stable persistent activations in a dominant population which is robust against interference. Reprinted with permission from Durstewitz and Seamans (2008).

and D'Esposito 2011) and serve to inform the interaction between motivation and cognition more generally.

Finally, other models—from more detailed implementations to more algorithmic approaches—have suggested that in addition to direct DA input to PFC, the basal ganglia (BG) can act as a gating mechanism by disinhibiting thalamocortical input to selective PFC subregions, allowing more refined and selective control of working memory updating (Frank et al. 2001; Todd et al. 2009). Again, these models have framed and guided a variety of subsequent findings showing complementary roles of PFC and BG in working memory updating versus maintenance, and have facilitated a computational theory of the role of the BG that extends beyond its classical role in motor control. This framework and that of dopaminergic signaling within PFC reviewed above provide clear translational implications for patients with mental illness (e.g., in attention-deficit/hyperactivity disorder, schizophrenia and, more generally, other frontostriatal disorders), by providing a coherent set of mechanisms relevant for understanding distractibility, attentional focus, and the interactions between reward and cognition.

In sum, this example provides a target set of phenotypes to study in mental illness: the trade-off of flexibility versus stability inherent in PFC-BG networks, and the dynamic modulation of this trade-off by DA inputs to both PFC and BG according to task objectives and reward maximization. Incorporating other neuromodulators into these theories is an important line of work with some promise; for example, reduced serotonin function in orbitofrontal cortex has been related to getting "stuck" in attractor states and associated with obsessions (Maia and Cano-Colino 2015).

Reinforcement Learning and Motivated Choice in Corticostriatal Circuits

One of the most seminal contributions of computational work in understanding systems and cognitive neuroscience was the proposal that phasic activity in midbrain DA neurons signal reward prediction errors (RPEs) (Figure 6.2a), with increases in activity for positive RPEs (outcomes that are better than expected) and dips below baseline for negative RPEs (worse than expected) (Montague et al. 1996). This model found striking correspondence between dopaminergic patterns of activity during simple reward conditioning tasks and RPEs as reflected in the temporal difference model of reinforcement learning, which allows an agent to learn precise expected reward values of various states in the environment, and—when augmented to learn the value of its own actions as well—to take actions that can maximize its cumulative reward. The quantitative link between DA and RPEs (both positive and negative) has since received enormous degree of support across species and methods (Schultz 2013), largely overturning older theories about the roles of DA in motor function and/

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

or reward signaling per se. Moreover, subsequent rodent genetic engineering studies have confirmed the causal importance of dopaminergic RPEs for inducing both Pavlovian and instrumental learning in ways that conform to learning theory. Human studies show neural markers of RPEs in striatal BOLD signals which are amplified by DA manipulations and correlate with reward learning (Pessiglione et al. 2006; Jocham et al. 2011).

Early theories also proposed that the downstream mechanism by which these DA signals promoted learning involved modification of corticostriatal synaptic strengths (e.g., Doya 2000; see also Figure 6.2b). Subsequent models expanded this notion by examining how the biology of this system supported the existence of two opponent systems that differentially learn from positive and negative RPEs (i.e., when outcomes are better and worse than expected), as a function of differential DA modulation of D1- and D2-containing medium spiny neurons, which act to promote action selection and avoidance (Frank 2005; Figure 6.2c). This model was motivated by decades of systems neuroscience including electrophysiological, pharmacological, and behavioral data. It suggests differential roles of these pathways, but was developed as an attempt to explain data from human Parkinson patients, whereby dopaminergic drugs can sometimes impair and sometimes enhance cognitive function. Many studies across species, over the last decade, have provided support for the basic model mechanisms, showing that modulation of D1 and D2 corticostriatal pathways is both necessary and sufficient for inducing reward/approach and aversive/avoidance learning, respectively (Hikida et al. 2010; Kravitz et al. 2012). Incorporating this opponent process into a refined algorithmic reinforcement learning facilitated a formal analysis of its properties, allowing for quantitative fits to multiple datasets, and provided a normative account to explain why this system might have evolved in this manner (Collins and Frank 2014; Figure 6.2d). It also provided an explanation for the finding that antipsychotics (and indeed striatal DA denervation, more generally) can induce an aberrant learning process resulting in progressive motor deterioration, beyond the direct effects of DA depletion on motor performance. Thus it hinted at a different mechanism for potential therapeutics. This same model has been applied to explain differential sensitivity to positive versus negative decision outcomes (a different sort of trade-off) across a range of conditions induced by dopaminergic dysregulation, including Tourette syndrome, schizophrenia, attention-deficit/hyperactivity disorder, pathological gambling, and substance abuse (Maia and Frank 2011).

Various extensions of simple reinforcement-learning models have also been developed and relate to underlying biology. First, basic models often assume a fixed *learning rate*; that is, the degree to which RPEs are used to update action value estimates and hence behavioral adjustment. More sophisticated models show how this learning rate can itself be dynamically adjusted to take into account different forms of uncertainty, so as to integrate optimally the informativeness of the incoming RPE relative to current knowledge and

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

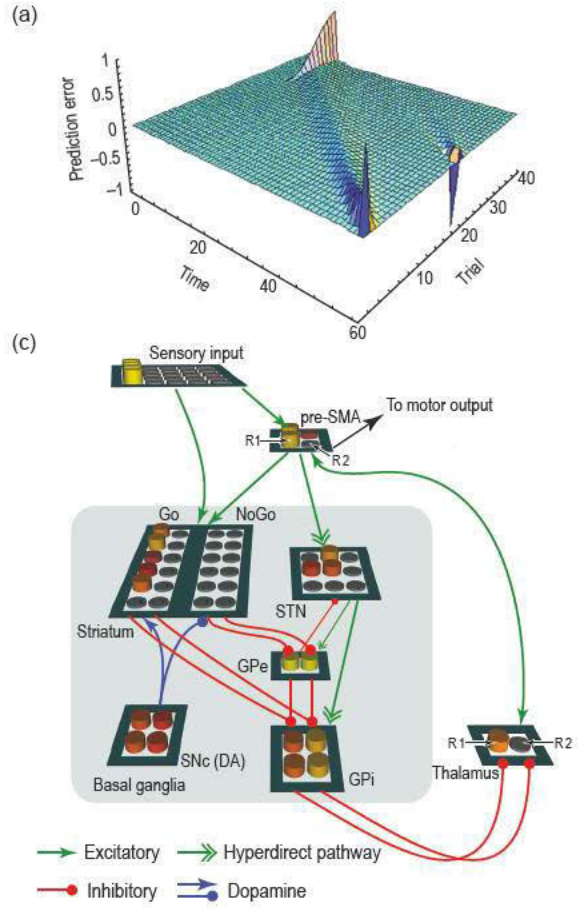

Figure 6.2 Dopamine, striatum, and reinforcement learning. (a) Temporal difference reinforcement-learning model showing phasic reward prediction error (RPE) signals initially when reward is delivered (end of the trial, final time step on *x*-axis) as well as with learning across trials (*z*-axis), where there is a lack of response to the reward itself; this prediction error signal is propagated back to the earliest predictor or reward (conditioned stimulus). On trial 20, the expected reward was withheld and a negative prediction error is observed. This pattern closely matches the phasic responses of midbrain dopamine (DA) neurons. Reprinted with permission from Schultz et al. (1997). (b) Models formalize how these DA RPEs are used to adjust the predicted values of sensory states V(s) and state-action pairs Q(s, a) in the striatum, with action selection dictated by comparison of action values in the downstream pallidum and subsequent disinhibition of the thalamocortical neurons coding for the most rewarding action. Reprinted with permission from Doya (2000). (c) Refined neural network of corticobasal ganglia circuit, with differential "Go" and "NoGo" striatal populations, representing positive and negative action values for given actions in pre-supplementary motor area (pre-SMA) and the current sensory state. Action selection is again governed by disinhibition (gating) of the corresponding column of thalamus, but where "NoGo" units provide evidence against a given action to prevent this disinhibition via the indirect pathway.

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

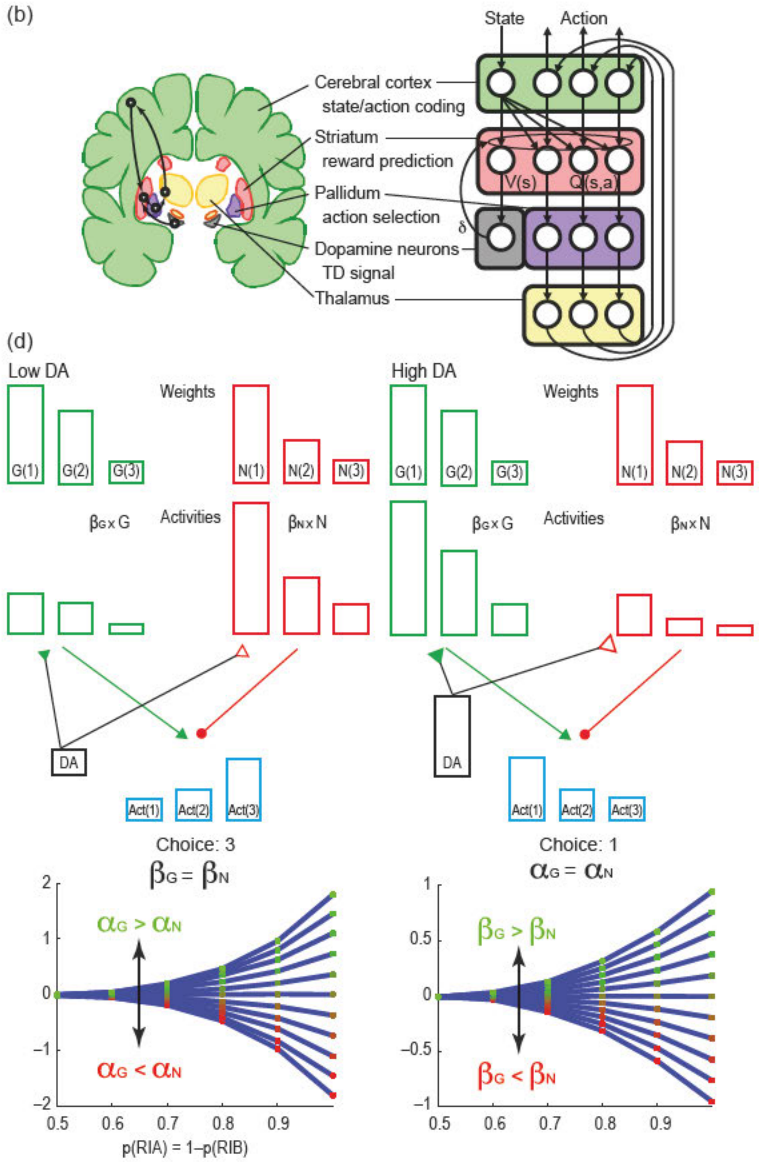

Figure 6.2 (continued) In the substantia nigra pars compacta (SNc), DA modulates excitability of "Go/NoGo" units via simulated D1 and D2 receptors, and phasic changes during RPEs drive opponent plasticity signals. The hyperdirect pathway from pre-SMA to subthalamic nucleus (STN) to globus pallidus internal (GPi) modulates the overall gating threshold by globally exciting GPi and making it more difficult for striatal "Go" signals to disinhibit the thalamus, thereby regulating impulsive choice. Individual neurons are cylinders, with instantaneous spike rate reflected by height and color. (*continued on next page*)

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

Figure 6.2 (continued) (d) Top: opponent actor learning (OpAL) model summarizing the core learning/choice computations of the neural network in algorithmic form, and capturing DA effects on both learning and incentive choice (Collins and Frank 2014). Separate G and N weights reflect learned propensities for each action to yield a positive or negative RPE. In the example, there are three actions and the corresponding learned G and N weights are shown in the top row. The middle row shows activity levels where DA levels during choice can be used to differentially amplify G or N weights via differential effects on excitability. In the "Low DA" condition, G weights are de-emphasized while differences among N weights are enhanced. Choice is governed by differences in activity levels; here the third choice which has the lowest cost is executed. In the "High DA" condition, for the same learned weights, the benefits are differentially amplified and Choice 1 is executed. Bottom: The OpAL model allows for asymmetry in the effects of DA on G versus N learning via the α parameters, and for asymmetry on their expression during choice, via the β parameters. Asymmetry in either set of parameters can produce differential sensitivity to probability of a choice leading to positive versus negative outcomes, and more so as the positive/negative probabilities become more deterministic.

to incorporate one's estimation of environmental volatility. Some evidence implicates the anterior cingulate cortex in tracking such volatility and adjusting learning rates (Behrens et al. 2007), while cholinergic interneurons may serve this function in the striatum (Franklin and Frank 2015). Recent studies suggest that this process is altered in individuals with high trait anxiety (Browning et al. 2015).

Second, rather than learn only values of simple stimulus-action pairs, *hierarchical reinforcement-learning* frameworks allow an agent to learn the values of, and select among, more abstract actions which themselves might involve multiple temporally extended sets of primitive actions (Botvinick et al. 2009). Once an abstract action is selected it can then carry out the sequence of state-action pairs that define it, allowing for more efficient reuse of previously learned actions (subgoals) that can be applied in the service of new goals. This hierarchical nesting of action selection and learning has been related to anatomical hierarchically nested rostral to caudal corticostriatal circuits, where anterior and lateral frontal circuits select actions which can then constrain the selection of lower-order actions in posterior loops.

Motivational Vigor and Incentive Choice

Although the learning theory of DA has been successful, some researchers prefer to emphasize the motivational aspects of DA in directly driving changes in vigor (speed with which actions are selected) and incentive choice (risky decision making). Computational models have simulated such effects as well by differential modulation of tonic (as opposed to phasic) DA, proposed to reflect the opportunity cost that would result from inaction (Niv et al. 2007). The opponent D1/D2 model ties together these DA roles on learning and choice: indeed, the mechanism by which DA modulates learning in this model is by

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

altering the excitability of these populations, and hence also affects their relative expression (activity levels) at the time of choice, providing a mechanism to dynamically modulate the emphasis on costs versus benefits of alternative choice strategies (Collins and Frank 2014; Figure 6.2d). Optogenetic studies show that effective action values can be enhanced or suppressed for particular choices when stimulating D1- or D2-expressing striatal neurons, respectively. However, this modeling work has also further highlighted an issue already recognized by many in the animal-learning community; namely, that many findings in reinforcement-learning experiments which appear to result from differential modulation of learning could instead reflect differential modulation of incentive choice, or vice versa. Thus careful designs are needed to tease apart their differential contributions.

Pavlovian factors can also affect instrumental performance. Pavlovian-to-instrumental transfer is the phenomenon by which stimuli taking on Pavlovian values can invigorate or inhibit instrumental action (Liljeholm and O'Doherty 2012). Computational models have quantified these effects and how they interact with instrumental learning (Huys et al. 2011a), and have further suggested that they involve both dopaminergic and serotonergic components as well as ventral striatal value-based modulation of dorsal striatal action (Boureau and Dayan 2011).

Model-Based Learning

Lacking in all of the above discussion on learning is a consideration of actions that are "goal-directed" (i.e., taken with the purpose of achieving a particular outcome), which often involves planning and forward thinking. Indeed, the DA–RPE hypothesis belongs to a special class of reinforcement-learning algorithms referred to as *model-free* in the sense that it involves learning incremental associative values (whether positive, negative, or combined), reflecting the statistical probability that an action will result in a good or bad outcome—a sort of net "gut-level" value—often intended to explain habits rather than goal-directed behaviors (Daw et al. 2005; Liljeholm and O'Doherty 2012). In contrast, a *model-based* learner will represent the expected outcomes of their actions using a cognitive map of the environment (Figure 6.3). The outcomes of actions could be other states that have no intrinsic value in and of themselves but open up yet other potential actions and consequent states. The model-based agent then conducts a mental search using their cognitive map to decide which action to take based on their goals. A model-based agent is much more flexible in that it can plan which course of action to take based on its current valuation of particular states and shift the course of action when those values change, without having to reexperience RPEs and incrementally adjust values for all relevant state-action pairs. However, it is also much more computationally demanding and time consuming, requiring the existence of a model of state transitions and the ability to search through the future trajectories while planning,

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

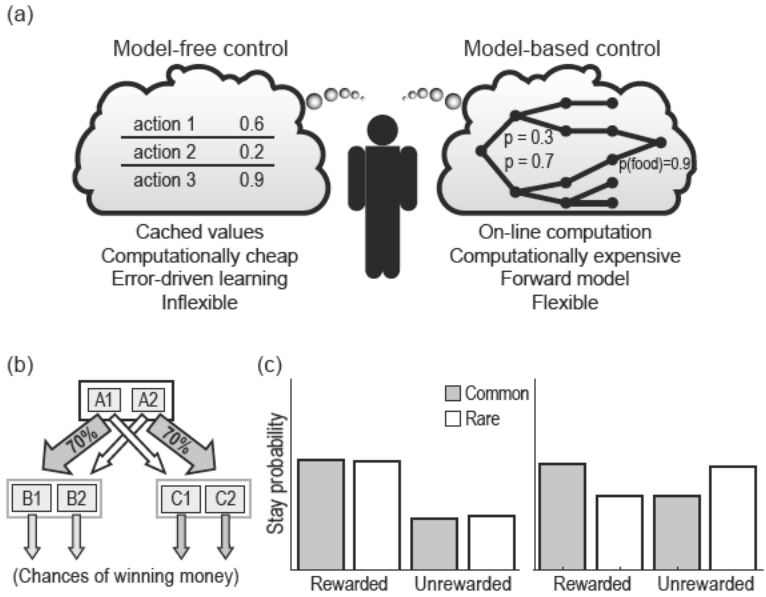

Figure 6.3 Model-based versus model-free reinforcement learning. (a) Model-free reinforcement learning as supported by dopaminergic RPEs allows a decision maker to learn a scalar value assigned to each action based on past reinforcement history; choices in this framework require only comparing such values without imagining their consequences for future choices. Such a system is efficient but inflexible when reinforcement contingencies of future states are altered and is typically invoked to explain habitual behaviors. In contrast, a model-based system reflects the predicted outcomes (subsequent states) that result from each action using a mental model of the environment, allowing a learner to make choices by searching this model and planning. Such a system is computationally demanding but flexible. Reprinted with permission from Smittenaar (2015). (b) Two-step task used to assess model-based versus model-free learning. Subjects choose between A1 and A2 which then yield subsequent B or C states. Choices at the bottom level yield different chances of obtaining rewards (money). Choice A1 results most commonly in the B state whereas A2 results in the C state, but rare transitions from A1 to C and A2 to B are also present. (c) Predicted probability of repeating the same top-level choice (A1 or A2) after rewards versus non-rewards, as a function of transition type from top to second level. A model-free agent will increase probability of staying (repeating the top-level choice) after rewards relative to non-rewards, regardless of the transition (it has no model of the transition structure). In contrast, a model-based agent shows an interaction: if rewarded on a rare transition, it knows that it will be more likely to reach the highly rewarding outcome on the next trial if it instead takes the top-level choice that most commonly transitions to that state. Reprinted with permission from Doll et al. (2012). Humans typically show a combination of model-based and model-free contributions, supported by interactive neural systems (see text).

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

thus involving working memory (and associated capacity limitations) and representations of the spatial environment. Indeed, rodent studies have shown striking evidence for forward projections of future states (in the hippocampus) and their values (in ventral striatum) as animals engage in decision making (van der Meer et al. 2010).

This example illustrates another key trade-off: the model-free system is efficient in allowing an agent to make choices rapidly based on accumulated values but is inflexible, whereas the flexible model-based system can allow for much more sophisticated choice strategies but is resource intensive. This trade-off suggests that the brain evolved to include both components, with various proposals debated about how they may compete or collaborate (Daw et al. 2005, 2011; Collins and Frank 2012). In rodents, model-based and model-free systems appear to depend on dissociable corticostriatal circuits, but in humans there is evidence for both competition and collaboration (Doll et al. 2012).

Thus multiple mechanisms are engaged in model-based reinforcement learning, including prefrontal-dependent working memory processes, hippocampal-dependent episodic and spatial memories, and an "arbitrator" for deciding which system should govern choice based on uncertainty or reliability (sometimes proposed to involve anterior cingulate cortex). Furthermore, the tendency to engage in model-based processing might itself be subject to motivational influences in addition to intrinsic capacities, and thus ventral striatum and anterior cingulate cortex modulation of cost-benefit trade-offs are equally relevant.

The interactions between model-based and model-free processes are only beginning to be uncovered. For example, the model-based system might train the model-free system so that it can engrain useful stimulus-response policies such that they no longer depend on cognitive resources. Reciprocally, model-free processes can be used to learn when to engage model-based systems. Model-based systems can also be used to learn representations of task structure—the variables in a task that matter, hidden causal states that govern contingencies—which can dramatically enhance the efficiency of model-free learning by collapsing across irrelevant features and facilitating generalization (Collins and Frank 2013; Wilson et al. 2014). Moreover, the complexity of model-based processing suggests that even when it is engaged, some shortcuts are often needed to prevent one from having to consider all possible courses of action; one such shortcut is a Pavlovian effect which drives subjects to avoid considering routes that elicit immediate negative states, even when this is suboptimal (Huys et al. 2012).

All of these mechanisms are ripe for further investigation into their disturbances in mental illness. Indeed, preliminary evidence suggests that a range of compulsive disorders are associated with reduced model-based processing (Voon et al. 2015). Further studies are, however, needed to examine the precise nature of these effects and their potentially dissociable underlying mechanisms. For example, model-based decision making can be impaired due to (a)

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

impaired *learning* of the model (i.e., impairments in detecting sequential transitions that describe the environment), (b) reduced tendency to *use* the model when making choices (reluctance to engage in deliberative processing required for planning), and/or (c) reduced motivation to engage in either model-based learning or model-based choice depending on the motivational stakes, cognitive load, etc.

Exploration versus Exploitation

To optimize learning, it is not always best to take the action that has higher expected value based on previous reward histories. Indeed, one should sometimes explore actions that have potential to provide yet better outcomes than the status quo (Figure 6.4). There are two main strategies that have been studied for balancing this exploration–exploitation trade-off. The first is to simply add some noise into the choice function: rather than deterministically choosing the options with highest reward values, a typical reinforcement-learning agent will make choices stochastically, allowing it to explore the values of actions it does not know. The most common of such choice functions is called *softmax* and is a logistic function that effectively adds more stochasticity to choices when the perceived values are more similar to each other, with more deterministic exploitation for values that are further apart. Such an algorithm is relatively simple to implement, with various proposals suggesting that cortical norepinephrine can dynamically modulate the noise in the choice function, and is itself regulated by recent task performance—encouraging more exploration during periods of poorer performance (Cohen et al. 2007). However, exploration might also demand cognitive control and prefrontal resources for overriding the dominant striatal tendency to exploit (Daw et al. 2006).

The second, a more strategic model-based approach to exploration, is to direct exploration toward those actions that have the greatest potential to be informative about the value of the current policy. Behavioral studies have shown that humans use a combination of both random and directed exploration (Wilson et al. 2014; Figure 6.4b). Further, fMRI and EEG studies have shown that the degree to which humans engage in directed exploration toward uncertain options is accompanied by rostrolateral PFC activity, which dynamically tracks the potential gain in information (relative uncertainty) that would result from such exploration (Badre et al. 2012); genetic and pharmacological studies suggest, however, that this tendency is modulated by prefrontal catecholaminergic function (Kayser et al. 2015). Notably, deficits in such uncertainty-driven exploratory behavior are correlated with anhedonia in patients with schizophrenia (Strauss et al. 2011). This finding might imply that anhedonia is not actually related to hedonics (i.e., an inability to experience pleasure)—indeed much evidence in the schizophrenia literature rejects that notion—but could

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

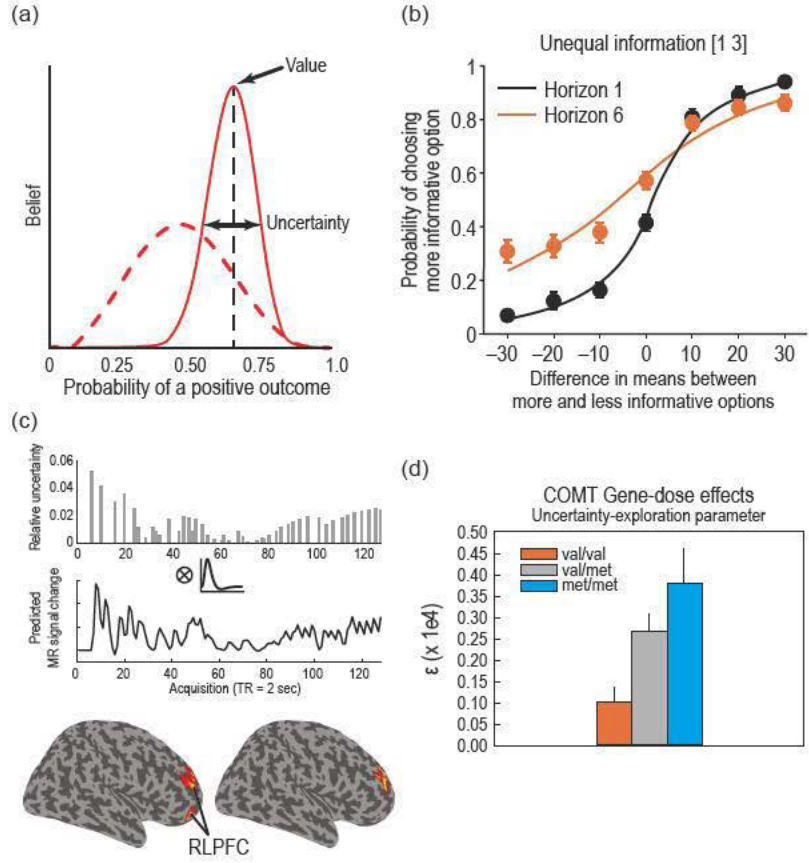

Figure 6.4 Exploration versus exploitation. (a) Adaptive reward-based choice requires not only comparing the values among alternative choices but considering the uncertainty in what those values are. Predicted rewards (or here, probabilities of obtaining a positive outcome) can be represented as entire belief distributions rather than single values. The solid curve represents an action that has high belief associated with value of 0.7 but other nearby values have similar belief levels. The dashed red line shows the belief distribution of an alternative action which has lower mean expected value but higher uncertainty (given limited experience), i.e., there is some possibility that its true value lies higher than that for the other action. Directed exploration thus takes informative actions that reduce this uncertainty. (b) Probability of choice increases as a function of mean difference in expected value based on limited samples, but there is a bias toward choice of the more informative action, particularly when the subject knows they will have the opportunity to make more choices (horizon 6) to capitalize on this information. Adapted with permission from Wilson et al. (2014). (c) fMRI study regressing trial-to-trial measures of information gain against BOLD activity reveals rostrolateral prefrontal cortex (RLPFC) region that tracks relative uncertainty about action values, and more so in explorers that use directed exploration behaviorally. Reprinted with permission from Badre et al. (2012). (d) Candidate gene affecting prefrontal catecholaminergic function modulates degree of directed exploration (Frank et al. 2009).

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

instead result from a reduced tendency to engage in those activities that could potentially improve their long-term situation.

Decision-Making Dynamics: From Simple Choice to Inhibitory Control

Aberrant decision making can also arise from changes to the decision process itself. Above we have been assuming that choices are made using some sort of comparison process among the learned reward values. While the reinforcement-learning literature focuses on how these values are acquired, the decision-making literature focuses on how choices are made within a given trial in the face of competing sources of evidence for each alternative, which may fluctuate based on momentary changes in perception, attention, or memory. Again, multiple levels of modeling have been applied to understand the decision-making process. One of the more popular frameworks is the drift diffusion model (Figure 6.5), which has been widely used for several decades in mathematical psychology; it accounts not only for choice proportions (which choices are made, given differences in the evidence for each option), but also for the full distributions of response times of those choices (Ratcliff and McKoon 2008). Sources of evidence in this framework can be perceptual (e.g., make a choice to discriminate whether you see an animal or a man-made object on a screen with different levels of discrimination difficulty, contrast and/or distractors), memory (e.g., determine whether an object presented has been studied before, given different levels of encoding), or based on reward values having multiple attributes (such as taste vs. health, e.g., choose among an apple or a cookie). In all cases, the drift diffusion model can be used to extract decision parameters that govern the choice process. The most relevant here are the "drift rate," which quantifies the amount of evidence inherent in the stimulus itself (or in the neural representation thereof), and the "decision threshold," which reflects the degree of evidence in favor of one option over the other before a participant is willing to commit to a choice. (It also has a Bayesian interpretation: given the stimulus presented, it reflects the likelihood ratio in favor of one option relative to the other; Gold and Shadlen 2007.) While changes to the drift rate or the decision threshold can produce changes in response times or choice proportions, they can be disentangled by examining simultaneously the choice proportions and response time distributions: small drift rates imply slower and more variable choices, higher decision thresholds lead to slower but more consistent (and accurate) choices. Any prior bias to select one option over the other (perhaps due to differential expectations before stimuli are observed) can be captured by the starting point, or bias parameter. More refined patterns of choice data and response time distributions can also be captured by estimating the degree of cross-trial

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

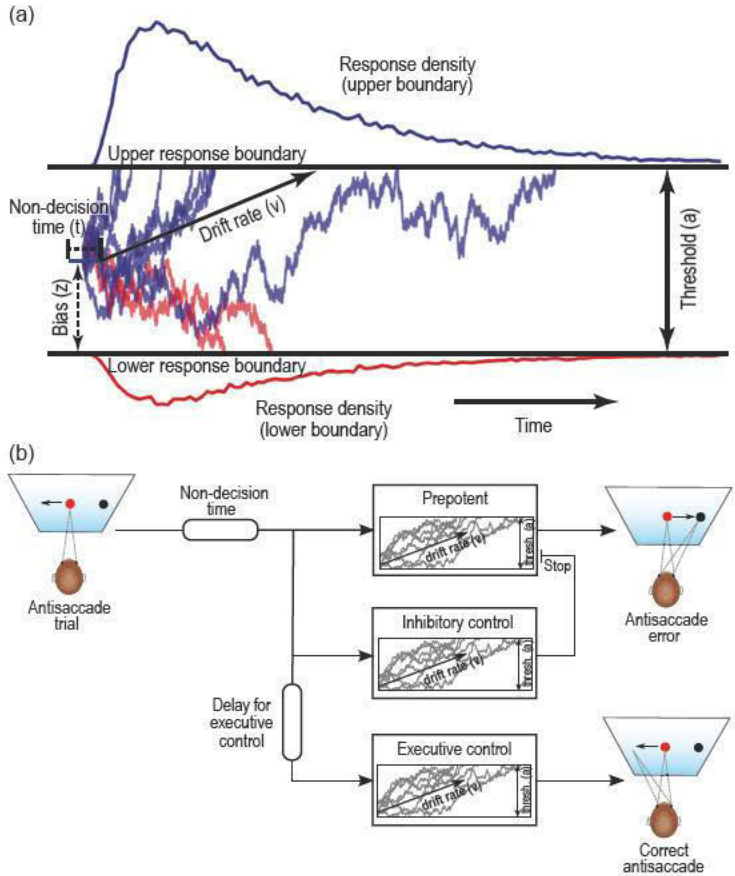

Figure 6.5 Decision-making dynamics. (a) The drift diffusion model is one of a larger class of sequential sampling models of simple decision making that can quantitatively capture choices and their response time distributions. It is commonly used to study neural mechanisms of the underlying process. Evidence for one choice over the other is noisily accumulated over time (*x*-axis) with average drift-rate *v* (reflecting the degree of evidence for one alternative over the other) until one of two boundaries (separated by threshold *a*) is crossed and a response is initiated. Trajectories of multiple drift processes (blue and red lines) are shown to illustrate the variability of this process across trials. The resulting response time distributions are shown for each choice. The non-decision time *t* accounts for time prior to (perception) and after (motor execution) the decision process; the bias (*z*) accounts for any predisposition toward one choice or the other. (b) Accumulator models show how different sources of evidence can accumulate independently to govern choices. This example shows how such models can be used to simulate dynamic processing within a trial supportive of inhibitory control, as in the antisaccade task, where an initial "intuitive" prepotent response drives accumulation toward one response boundary, an inhibitory signal accumulates to stop this process, and later information derived from cognitive rules can implement executive control to modify responding toward the correct action. Adapted with permission from Noorani and Carpenter (2012).

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

variability in these parameters (drift rate and starting point), which can reflect attentional lapses and/or changes in neural noise.

Beyond providing an abstract formalism that would allow psychologists to quantify distinct cognitive processes governing simple choice, many neuroscientific investigations have shown evidence for the otherwise unobservable internal processes of such models. Much of this evidence comes from perceptual decision-making tasks in which electrophysiological signals (e.g., spike rates in parietal cortex and striatum) exhibit the predicted ramping of decision variables quantifying the degree of evidence (or expected reward value) for one option over the other, with the slope of this ramping proportional to the signal-to-noise ratio in the stimulus; choices are executed once these neural signals reach a critical threshold (Gold and Shadlen 2007; Ratcliff and McKoon 2008). In reward-based decision making, the drift rate is proportional to relative difference in reward values among alternative options and is modulated by the subject's visual attention and reflected in striatum and ventromedial PFC (Lim et al. 2011). Moreover, choice values reflect multiple attributes: those involving impulsive urges (e.g., unhealthy foods) are encoded earlier in ventromedial PFC and are subsequently modified by long-term goals (Harris et al. 2013); competition among each attribute can occur at various hierarchical stages (Hunt et al. 2014). These models can also be modified to simulate decisions in inhibitory control such as the antisaccade task, where a prepotent response is elicited by an imperative stimulus, but where cognitive rule-based representations (e.g., in PFC) can modify the decision process later in time (Noorani and Carpenter 2012). More generally, these models can be used to estimate the separable contributions of mechanisms giving rise to fast automatic response tendencies, those associated with inhibiting such responses, and those associated with guiding behaviors toward controlled responses.

Mechanistic neural models also articulate how these different choice processes are implemented in cortical attractor networks (Wang 2012) and in the BG (Lo and Wang 2006; Bogacz and Gurney 2007; Ratcliff and Frank 2012). One critical variable that should be evident in terms of aberrant decision making is the decision threshold: if, as noted above, impulsive urges contribute to value signals early during the choice process, then a mechanism to regulate decision thresholds is critical to regulate decision making. Indeed, mechanistic neural models have specified how the decision threshold is malleable and subject to communication between frontal cortex and BG. In particular, when subjects experience "decision conflict" (i.e., when there are multiple decision alternatives or when the execution of a rule-guided action requires overriding an initial prepotent response), the frontal cortex signals this information to the subthalamic nucleus (STN) (see Figure 6.2c) which—by effectively raising the decision threshold—makes it more difficult for initial striatal valuation signals to impulsively govern choice. This role for the fronto-STN pathway in decision-threshold regulation and inhibitory control is corroborated by

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

neuroimaging, intracranial electrophysiology, and deep brain stimulation studies (e.g., Hikosaka and Isoda 2010; Cavanagh et al. 2011). Indeed, deep brain stimulation of STN can induce impulsivity in patients' daily lives and prevent them from adjusting decision thresholds as a function of conflict. These studies provide an opportunity for developing phenotypes that identify when such mechanisms are aberrant and likely to be causing dysfunction. Moreover, this same STN mechanism could play a role not only in impulsivity (linked to a reduced decision threshold) but in the opposite scenario (when STN and/or its cortical inputs are overactive, linked to too high a decision threshold). It might also be implicated in perfectionism, obsessiveness, and deficits in proactive reward-based decisions.

This last example highlights that a particular phenomenon of impulse control disorders can arise from multiple mechanisms. Here, though, the emphasis is on the need to regulate the decision threshold to consider alternative goals given an impulsive urge, whereas earlier it was on the imbalance in the sensitivity to prospective positive versus negative decision outcomes due to changes in DA function. Indeed, the modeling framework and relevant task paradigms have facilitated the ability to dissect components of impulsivity related to insensitivity to adverse consequences (which are affected by DA medications given to patients with Parkinson disease and associated with pathological gambling), from those involved in disinhibition during conflict-based decision making that is affected by deep brain stimulation of STN (Frank et al. 2007b). Moreover, yet other forms of impulsivity may involve differential discounting of long-term versus immediate rewards (e.g., McClure et al. 2004), the neural mechanisms of which are also intensely studied.

Social Interactions

Although we have focused in this chapter on simple decision-making and learning tasks, computational cognitive neuroscience approaches have also examined social interactions. For example, game theory, traditionally developed in economics, has been recently applied to understand how humans interact with each other in various cooperative and competitive environments that require theory of mind and value representation of self versus other (Montague et al. 2012). This research program shows how computational approaches can be used to infer latent processes involved in social decision making, much as latent processes are involved in inferring task variables in complex reinforcement-learning environments alluded to above. For example, I may view the value of a particular choice in terms of its immediate outcome, but if that outcome also depends on another person's goals and intents, we can develop a model that incorporates their beliefs and goals and determine if that should influence our own decisions and associated values. This type of approach is

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.

beginning to see some useful application to explore how these processes are altered in disorders such as autism.

Conclusion

This selective and quite incomplete overview covers only a small portion of computational approaches to a restricted set of domains within the larger field of cognitive neuroscience. Nevertheless, it has highlighted how computational models at multiple levels of description have contributed to a richer understanding of the neural basis of cognitive function, including several examples for how these have been or could be capitalized to understand the failure modes of such functions in mental illness. Much more work is needed to study which individual mechanisms can be assessed using refined cognitive tasks, neural measures, and quantitative models, as well as how these mechanisms interact to form distinct functional profiles (Stephan and Mathys 2014; Wiecki et al. 2015).

From "Computational Psychiatry: New Perspectives on Mental Illness,"
A. David Redish and Joshua A. Gordon, eds. 2016. Strüngmann Forum Reports, vol. 20,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03542-2.