Schizophrenia: A Computational Reinforcement Learning Perspective

Michael J. Frank

Department of Psychology, Program in Neuroscience, University of Arizona, 1503 E University Boulevard, Tucson, AZ 85721

As one of the most complex neurocognitive disorders, schizophrenia (SZ) is a devastating condition for which the underlying sources are far from being fully understood. Indeed, it is likely that there are multiple etiologies to the disease and heterogeneity within the population. Moreover, it is impossible to understand from a purely mechanistic basis how a patient would come to believe so strongly in delusions as to, for example, gouge out his own eyes.

Nevertheless, science marches forward, and the last 30 years or so have produced a wealth of knowledge regarding some of the risk factors, genetics, pharmacology, cognitive deficits, and underlying neurobiology associated with the disease. In part because of the efficacy of antipsychotic treatments via dopamine D2 receptor blockade, the majority of this research focuses on dysfunctions of the dopaminergic system, in both frontal cortex and basal ganglia, thought to be related to negative and positive symptoms, respectively. At the neurocognitive level, much of the focus has been on dysfunction within dorsolateral prefrontal cortical circuits and their contributions to working memory, cognitive control, and attentional shifting. While dopamine plays a critical role in all these processes, it is perhaps more centrally related to aspects of motivational processing, which is surprisingly understudied in SZ. Indeed, it is possible to account for many of the frontal-dependent cognitive deficits in SZ by positing a more core deficit in the motivational “gating” system for determining which information patients should “care” about and what they might ignore.

Given the complexity of neural circuits involved in both cognitive and motivational functions, it becomes dauntingly difficult to capture the possible interactions of these circuits, and particularly how they are disrupted in SZ, with simple verbal depictions and static anatomical diagrams. Here I consider the potential application of computational neural network models as a principled and dynamic tool for exploring these interactions and psychopathology associated with dopaminergic dysfunction in SZ and which can lead to new testable predictions at both the neural and behavioral levels. These models enable one to simulate various anatomical and physiological pieces of data, using mathematical equations that capture how groups of neurons communicate activity to other neurons within and between brain areas. By incorporating aspects of neuronal physiology, connectivity, and synaptic plasticity within the basal ganglia–frontal cortical system, one can examine dynamics of this circuitry and how it may go awry. At the same time, it is not tractable to try to incorporate every known biological detail into a model, particularly when the goal is to discover how an entire system of brain regions interact to produce behavior. Thus, the models are also constrained by the need to account for existing data at these higher levels, such as effects of focal lesions or pharmacological manipulation on behavior. Critically, the models make new predictions about how the system works that would likely not have emerged otherwise and often were not conceived by the modeler prior to being built. Models can then be tested and refined and their implications explored in neurological conditions.

To sum up a large body of basic research, models of frontostratial function have generally suggested that these circuits support the following—(1) action selection: as in when making a choice among multiple competing alternatives and (2) reinforcement learning: as in modifying expectations and behavior following positive and negative outcomes. For the former process, “actions” to be selected include both lower level motor programs, consistent with the traditionally ascribed role of the basal ganglia in motor control, and higher level cognitive actions, such as when and when not to update/manipulate the contents of working memory. Reinforcement learning then operates on these actions such that adaptive actions are more likely to be repeated, whereas maladaptive actions are suppressed. Critically, according to both the models and available electrophysiological evidence, positive outcomes are reflected in terms of deviations from current expectations, a term referred to as a “positive prediction error,” and are encoded by phasic bursts of dopamine. Similarly, negative prediction errors are encoded by phasic dips or pauses in dopaminergic activity. These phasic bursts and dips modify corticostratial synaptic plasticity, allowing the system

---

1 To whom correspondence should be addressed; tel: 520-626-4787, fax: 520-621-9306, e-mail: mfrank@u.arizona.edu.
to incrementally become more likely to produce actions that are adaptive and to avoid those that are maladaptive.

Importantly, these models have generated several testable predictions for clinical populations and pharmacological manipulations. In brief, cognitive experiments have provided much support for the idea that dopamine manipulation can affect reinforcement learning and motivational processes. For example, in probabilistic reinforcement learning paradigms, Parkinson’s disease patients, who have low striatal dopamine levels, are relatively worse at ‘Go learning’ from positive prediction errors resulting from their decisions than they are at ‘NoGo learning’ from negative prediction errors; this relationship reverses while they are on dopaminergic agonist medications as predicted by the models. In addition to SZ patients showing relatively selective deficits in probabilistic Go but not NoGo learning signals across time, they also showed profound reductions in the tendency to rapidly adapt choices on a trial-to-trial basis following a single instance of reinforcement feedback. These rapid adaptations are thought to rely on different cognitive and neural systems than those involved in integrating probabilities across time, potentially linked to orbitofrontal cortex rather than striatum. Supporting this interpretation, deficits in rapid adaption were correlated with negative symptoms, thought to stem from frontal cortical degradation, and patients with orbitofrontal damage show similarly slowed acquisition in analogous reinforcement tasks.

What are the implications of such theories and experiments for schizophrenia? Due to dopaminergic dysregulation in frontostriatal circuits, it is possible that thoughts and actions that would normally be suppressed are actually reinforced. This would be manifest in terms of both changes in cognitive performance, but perhaps more primarily in terms of the underlying motivation. By this account, delusions may be partially understood in terms of faulty prediction error signals that fail to discriminate between logical, rational, or adaptive associations such that patients would sometimes attend to internal or external stimuli that they should ignore, and ignore those that they should attend.

Indeed, functional neuroimaging studies reveal that prefrontal cortex is not always hypoactive, but sometimes hyperactive, in SZ consistent with a dysfunctional gating process. Imaging has also revealed that striatal reinforcement prediction error signals are disrupted both with psychosis and delusions. In the same probabilistic reinforcement learning paradigm previously used in Parkinson’s patients (see above), medicated SZ patients showed relatively impaired ‘Go learning’ from positive prediction errors, while showing spared ‘NoGo learning’ from negative prediction errors. Similarly, patients fail to show the normal implicit tendency to speed responses when faced with high reward incentives, a process known to depend on striatal dopamine. In our models, all the above Go learning deficits are accounted for by reduced striatal D1 receptor function, compounded by noisy phasic DA signals that do not appropriately report the strength of positive prediction errors. Further, the spared NoGo learning may be attributed to D2 receptor blockade by antipsychotic medications, which would actually potentiate synapses in the NoGo pathway, such that learning in medicated SZ patients is similar to that of nonmedicated PD patients. Interestingly, this same D2 mechanism in our computational model accounts for a variety effects resulting from haloperidol administration in rodents, which leads to a progressive sensitization of catalepsy expression that is context dependent (T. V. Wiecki, K. Riedinger, A. Meyerhofer, W. J. Schmidt, and M. J. Frank, unpublished data).

SZ has substantial genetic heritability, in the range of 80%. One approach to understand specific components of the disease is to focus on particular “intermediate phenotypes” that are related to specific genetic factors and which contribute to a subset of the disease (rather than the full-blown pathology associated with multiple neurobiological correlates). Given the focus on the dorsolateral prefrontal cortex and cognitive dysfunction in SZ, it should perhaps not be surprising that a similar focus has been applied in the genetic domain. The same logic can be applied to understanding individual differences in reinforcement learning, as informed by computational models, resulting from genetic factors controlling striatal and frontal dopaminergic function. Indeed, there are now multiple studies linking candidate genes associated with SZ and changes in reinforcement learning from both behavioral and functional neuroimaging. In particular, independent genes that control different aspects of striatal dopaminergic function have been associated with probabilistic “Go” and “NoGo” learning. The DARPP-32 protein is known to be dependent for striatal D1 receptor–mediated synaptic plasticity in response to rewarding events. A polymorphism within this gene alters striatal function in SZ (thus far shown in working memory tasks). In healthy individuals, this same polymorphism is predictive of probabilistic Go learning. In contrast, the DRD2 gene, which is associated with striatal D2 receptor density, is predictive of NoGo learning. Thus, it may be informative to study whether these genes partially determine antipsychotic effects on motivational changes in SZ. Finally, the well-studied COMT gene controlling prefrontal, but not striatal, dopaminergic function was only associated with rapid trial-to-trial adaptation but not probabilistic Go or NoGo learning. Thus, these dissociable reinforcement processes in healthy individuals suggests that these may also be related to heterogeneity within the SZ population.
In a noisy world with mixed reinforcement signals, how does one determine whether to respond based on the most recent outcomes or to continue to go with what they had learned over the course of their history? Similarly, a major question in computational reinforcement learning is how does an agent know when it is appropriate to “exploit” their current strategy which may work to a certain degree, and when should one strategically “explore” other alternatives because they might be even better? Recent neuroimaging data implicate prefrontal function for making these strategic exploratory decisions. Preliminary genetic data in our laboratory implicate the COMT gene in predicting individual differences in these kinds of exploratory decisions, which are also expected to be aberrant in SZ—and may provide a computational explanation for prefrontal-dependent negative symptoms of the disease.

By combining studies in animals, pharmacology, and genetics with theoretical models of dopaminergic function with frontostriatal circuits, the hope is to shed light on specific motivational processes that may go awry in SZ and how these may be altered for better or worse by antipsychotic medication. It must be emphasized that data from studies with SZ patients can neither confirm nor falsify basic mechanisms of the neurobiological models (which themselves are in need of further refinement). Nevertheless, computational models provide an explicit framework which, in concert with empirical research, can provide a valuable tool to understanding vexing and multivariate problems associated with this complex disorder.

Funding

NIMH grant R01 MH080066-01.

References

5. Seeman P. Dopamine receptors and the dopamine hypothesis of schizophrenia. Synapse. 1987;1:133–152.


46. Klein A, Schmidt WJ. Catalepsy intensifies context-dependently irrespective of whether it is induced by intermittent or chronic dopamine deficiency. *Behav Pharmacol*. 2003;14:49–53.


