Schizophrenia Bulletin vol. 34 no. 6 pp. 1008–1011, 2008
doi:10.1093/schbul/sbn123
Advance Access publication on September 12, 2008

Schizophrenia: A Computational Reinforcement Learning Perspective

Michael J. Frank[1]

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.[1–4] In part because of the efficacy of antipsychotic treatments via dopamine D2 receptor blockade,[5] 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.[6] 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.[7,8] 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.[9,10] 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.[11]

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.[12]

To sum up a large body of basic research, models of frontostriatal 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.[13–19] 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.[17,20–22] 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.[14,19,23,24] Similarly, negative prediction errors are encoded by phasic dips or pauses in dopaminergic activity. These phasic bursts and dips modify corticostriatal 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.[25–27] Similarly, the ability to know when to "gate" sensory information into working memory is dependent on intact basal ganglia and dopaminergic processes.[28–32]

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.[33,34]

Indeed, functional neuroimaging studies reveal that prefrontal cortex is not always hypoactive, but sometimes hyperactive, in SZ,[35,36] consistent with a dysfunctional gating process. Imaging has also revealed that striatal reinforcement prediction error signals are disrupted both with psychosis and delusions.[37,38] 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.[39] Similarly, patients fail to show the normal implicit tendency to speed responses when faced with high reward incentives,[40] a process known to depend on striatal dopamine.[41,42] 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,[43,44] such that learning in medicated SZ patients is similar to that of nonmedicated PD patients.[25] 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).[45,46]

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.[39] 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.[10,15] Supporting this interpretation, deficits in rapid adaption were correlated with negative symptoms,[39] thought to stem from frontal cortical degradation,[6] and patients with orbitofrontal damage show similarly slowed acquisition in analogous reinforcement tasks.[47,48]

SZ has substantial genetic heritability, in the range of 80%.[49] 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).[50] 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.[4] 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.[51–55]. 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.[56,57] A polymorphism within this gene alters striatal function in SZ (thus far shown in working memory tasks[4]). In healthy individuals, this same polymorphism is predictive of probabilistic Go learning.[52] In contrast, the DRD2 gene, which is associated with striatal D2 receptor density,[58] is predictive of NoGo learning.[52] 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[49,59] was only associated with rapid trial-to-trial adaptation but not probabilistic Go or NoGo learning.[52] 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.[60] 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

1. Goto Y, O'Donnell P. Altered prefrontal cortex-nucleus accumbens information processing in a developmental animal model of schizophrenia. *Ann N Y Acad Sci*. 2003;1003: 398–401.
2. Goto Y, Grace AA. The dopamine system and the pathophysiology of schizophrenia: a basic science perspective. *Int Rev Neurobiol*. 2007;78C:41–68.
3. Lodge DJ, Grace AA. Aberrant hippocampal activity underlies the dopamine dysregulation in an animal model of schizophrenia. *J Neurosci*. 2007;27:11424–11430.
4. Meyer-Lindenberg A, Straub RE, Lipska BK, et al. Genetic evidence implicating DARPP-32 in human frontostriatal structure, function, and cognition. *J Clin Invest*. 2007;117:672–682.
5. Seeman P. Dopamine receptors and the dopamine hypothesis of schizophrenia. *Synapse*. 1987;1:133–152.
6. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry*. 1987;44:660–669.
7. Cohen JD, Servan-Schreiber D. Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol Rev*. 1992;99:45–77.
8. Barch DM, Carter CS, Braver TS, et al. Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Arch Gen Psychiatry*. 2001;58: 280–288.
9. Barch DM. Emotion, motivation, and reward processing in schizophrenia spectrum disorders: what we know and where we need to go. *Schizophr Bull*. 2008;34:816–818.
10. Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA. Reward Processing in schizophrenia: a deficit in the representation of value. *Schizophr Bull*. 2008;34:835–847.
11. Braver TS, Barch DM, Cohen JD. Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biol Psychiatry*. 1999;46:312–328.
12. Frank MJ, Scheres A, Sherman SJ. Understanding decision making deficits in neurological conditions: insights from models of natural action selection. *Philos Trans R Soc Lond B*. 2007;362:1641–1654.
13. Houk JC, Adams JL, Barto AG. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk JC, Davis JL, Beiser DG, eds. *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press; 1995:233–248.
14. Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J Cogn Neurosci*. 2005;17:51–72.
15. Frank MJ, Claus ED. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev*. 2006;113:300–326.
16. Brown JW, Bullock D, Grossberg S. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Netw*. 2004;17:471–510.
17. Houk JC. Agents of the mind. *Biol Cybern*. 2005;92:427–437.
18. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol Cybern*. 2001;84:401–410.
19. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005;8:1704–1711.
20. Frank MJ, Loughry B, O'Reilly RC. Interactions between the frontal cortex and basal ganglia in working memory: a computational model. *Cogn Affect Behav Neurosci*. 2001;1: 137–160.
21. O'Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput*. 2006;18:283–328.
22. Gruber AJ, Dayan P, Gutkin BS, Solla SA. Dopamine modulation in the basal ganglia locks the gate to working memory. *J Comput Neurosci*. 2006;20:153–166.
23. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*. 1996;16:1936–1947.
24. Schultz W. Getting formal with dopamine and reward. *Neuron*. 2002;36:241–263.
25. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science*. 2004;306:1940–1943.
26. Frank MJ, Samanta J, Moustafa AA, Sherman SJ. Hold your horses: impulsivity, deep brain stimulation and medication in Parkinsonism. *Science*. 2007;318:1309–1312.
27. Cools R, Altamirano L, D'Esposito M. Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*. 2006;44:1663–1673.

28. McNab F, Klingberg T. Prefrontal cortex and basal ganglia control access to working memory. *Nat Neurosci*. 2008;11: 103–107.
29. Frank MJ, O'Reilly RC. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci*. 2006;120:497–517.
30. Frank MJ, Santamaria A, O'Reilly RC, Willcutt E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology*. 2007;32:1583–1599.
31. Cools R, Sheridan M, Jacobs E, D'Esposito M. Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *J Neurosci*. 2007;27:5506–5514.
32. Moustafa AA, Sherman SJ, Frank MJ. A dopaminergic basis for working memory, learning, and attentional shifting in Parkinson's Disease. *Neuropsychologia*. 2008;46:3144–3156.
33. Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry*. 2003;160:13–23.
34. Smith A, Li M, Becker S, Kapur S. Dopamine, prediction error and associative learning: a model-based account. *Network*. 2006;17:61–84.
35. Callicott JH, Mattay VS, Verchinski BA, Marenco S, Egan MF, Weinberger DR. Complexity of prefrontal cortical dysfunction in schizophrenia: more than up or down. *Am J Psychiatry*. 2003;160:2209–2215.
36. Manoach DS. Prefrontal cortex dysfunction during working memory performance in schizophrenia: reconciling discrepant findings. *Schizophr Res*. 2003;60:285–298.
37. Corlett PR, Murray GK, Honey GD, et al. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*. 2007;130:2387–2400.
38. Murray GK, Corlett PR, Clark L, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry*. 2008;13:239, 267–239, 276.
39. Waltz JA, Frank MJ, Robinson BM, Gold JM. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry*. 2007;62:756–764.
40. Murray GK, Clark L, Corlett PR, et al. Incentive motivation in first-episode psychosis: a behavioural study. *BMC Psychiatry*. 2008;8:34.
41. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning or incentive salience. *Brain Res Rev*. 1998;28:309–369.
42. Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. 2007;191:507–520.
43. Centonze D, Usiello A, Costa C, et al. Chronic haloperidol promotes corticostriatal long-term potentiation by targeting dopamine D2L receptors. *J Neurosci*. 2004;24:8214–8222.
44. Shen W, Flajolet M, Greengard P, Surmeier DJ. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*. 2008;321:848–851.
45. Amtage J, Schmidt WJ. Context-dependent catalepsy intensification is due to classical conditioning and sensitization. *Behav Pharmacol*. 2003;14:563–567.
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.
47. Chase HW, Clark L, Myers CE, et al. The role of the orbitofrontal cortex in human discrimination learning. *Neuropsychologia*. 2008;46:1326–1337.
48. Wheeler EZ, Fellows LK. The human ventromedial frontal lobe is critical for learning from negative feedback. *Brain*. 2008;131:1323–1331.
49. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry*. 2003;60:1187–1192.
50. Tan H-Y, Callicott JH, Weinberger DR. Intermediate phenotypes in schizophrenia genetics redux: is it a no brainer? *Mol Psychiatry*. 2008;13:233–238.
51. Cohen MX, Young J, Baek J-M, Kessler C, Ranganath C. Individual differences in extraversion and dopamine genetics predict neural reward responses. *Brain Res Cogn Brain Res*. 2005;25:851–861.
52. Frank MJ, Moustafa AA, Haughey H, Curran T, Hutchison K. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci USA*. 2007;104:16311–16316.
53. Klein TA, Neumann J, Reuter M, Hennig J, Cramon DY, Ullsperger M. Genetically determined differences in learning from errors. *Science*. 2007;318:1642–1645.
54. Forbes EE, Brown SM, Kimak M, Ferrell RE, Manuck SB, Hariri AR. Genetic variation in components of dopamine neurotransmission impacts ventral striatal reactivity associated with impulsivity. *Mol Psychiatry*. September 25, 2007. doi: 10.1038/sj.mp.4002086.
55. Yacubian J, Sommer T, Schroeder K, et al. Gene-gene interaction associated with neural reward sensitivity. *Proc Natl Acad Sci USA*. 2007;104:8125–8130.
56. Calabresi P, Gubellini P, Centonze D, et al. Dopamine and cAMP-regulated phosphoprotein 32 kDa controls both striatal long-term depression and long-term potentiation, opposing forms of synaptic plasticity. *J Neurosci*. 2000;20: 8443–8451.
57. Stipanovich A, Valjent E, Matamales M, et al. A phosphatase cascade by which rewarding stimuli control nucleosomal response. *Nature*. 2008;453:879–884.
58. Hirvonen M, Laakso K, Rinne JO, Pohjalainen T, Hietala J. C957T polymorphism of the dopamine D2 receptor (DRD2) gene affects striatal DRD2 availability in vivo (Corrigendum). *Mol Psychiatry*. 2005;10:889.
59. Slifstein M, Kolachana B, Simpson EH, et al. COMT genotype predicts cortical-limbic D1 receptor availability measured with [11C]NNC112 and PET. *Mol Psychiatry*. 2008;13: 821–827.
60. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature*. 2006;441:876–879.