Published in final edited form as:

Neuroreport. 2008 March 5; 19(4): 459–462.

A model of behavioral treatments for self-mutilation behavior in Lesch-Nyhan syndrome

Eric A. Zilli^{a,b,*} and Michael E. Hasselmo^{a,b}

a Program in Neuroscience, Boston University. Center for Memory and Brain, 2 Cummington St., Boston, Massachusetts 02215, U.S.A., (617) 353-1397, FAX: (617) 358-3269

b Department of Psychology, Boston University. Center for Memory and Brain, 2 Cummington St., Boston, Massachusetts 02215, U.S.A., (617) 353-1397, FAX: (617) 358-3269

Abstract

Lesch-Nyhan syndrome is a rare genetic disorder, caused by a mutation in the gene coding for the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT), which is characterized by hyperuricemia and its associated symptoms along with motor disorders and compulsive self-mutilation. We show that the temporal difference learning algorithm that has been often used to interpret dopaminergic activity in the basal ganglia offers an explanation for the self-mutilation behaviors. We propose that a dysfunctional dopamine signal inadvertently reinforces early, accidental injurious behavior that is initially caused by clumsiness due to the motor disorders. Simulations of this model reproduce findings on the results of behavioral treatments for dealing with self-mutilation behaviors.

Keywords

differential reinforcement; dopamine; punishment; reinforcement learning; self injury; time outs

Introduction

Lesch-Nyhan syndrome (LNS) is a rare, X-linked genetic disorder caused by a mutation in the gene coding for the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT) which is important in the metabolism of purines [1]. Along with the resulting hyperuricemia, LNS is accompanied by cognitive deficits and a number of motor disorders including choreoathetosis, spasticity, and hypertonia [2]. More striking, however, is the common behavior of compulsive, often severe self-mutilation that begins around the age of 2 years, and may include biting off parts of one's fingers, tongue, and lips, and other behaviors that produce harm to the patient and others [3], during which the patients feel normal pain [1]. Specific accounts relating the cause of these behaviors to the underlying physiology have not yet appeared, and pharmacological treatments of self-mutilation behavior with a variety of drugs produce inconsistent results [4].

Here we model behavioral treatments of these self-injurious behaviors by focusing on data showing that neuromodulators in the basal ganglia (BG), particularly dopamine (DA), are dysfunctional in the brains of LNS patients [4,5]. Neurochemical studies of LNS patients' brains show large reductions in DA levels [6–8], elevated numbers of DA receptors [8], and decreased levels of DA transporters [9]. Experimental work on the firing patterns of DA neurons in monkeys has recently converged with theoretical work on instrumental conditioning

*To whom correspondence should be addressed. E-mail: zilli@bu.edu.

using temporal difference (TD) learning algorithms [10], which provides a quantitative link between DA and learning and behavior [11–13].

Methods

We simulate dysfunction of the DA system using a TD algorithm that we summarize here [10,12]. The following three equations summarize the learning that would occur after the model takes an action a in state s which results in the new state s' and reward r .

$$\delta = r + V(s') - V(s) \quad (1)$$

$$V(s) \leftarrow V(s) + \alpha \delta \quad (2)$$

$$Q(s, a) \leftarrow Q(s, a) + \alpha \delta \quad (3)$$

where $0 \leq \alpha \leq 1$ is a learning rate parameter. This is simply the calculation of a prediction error (equation 1) followed by updating predictions of state values V and action values Q .

The agent selects actions according to a Boltzmann distribution. The probability of selecting action a in state s is:

$$\Pr(a|s) = \frac{\exp(Q(s, a))}{\sum_{a'} \exp(Q(s, a'))} \quad (4)$$

This is the basic actor-critic reinforcement learning framework that is used to model the healthy DA system in animal basal ganglia [10,12]. We simulate a number of modifications to this framework as described below, each of which consists of replacing one of these equations with a modified equation.

Results

Anderson and colleagues [14] summarized work on operant conditioning of LNS patients saying they are “either reinforced by, do not learn from, or are unable to inhibit a response associated with a ‘punishment’ contingency.” Specifically, they showed that punishing attempts at self-injury usually resulted in an increased frequency of attempts. In the terms of TD learning, it appears that patients are ultimately compelled to select the action with the lowest action value instead of the highest action value (e.g. inflicting pain on self or others, answering questions incorrectly [15]). This could correspond to replacing equation 4 with

$$\Pr(a|s) = \frac{\exp(-Q(s, a))}{\sum_{a'} \exp(-Q(s, a'))} \quad (4a)$$

However, the same behavior might be produced by changing the sign of one of the seven other terms in equations 1–4.

Simulations of the above equations were performed using two actions, one which always produced a reward of -1 (“bad” action) and one which always produced a reward of $+1$ (“good” action). Each simulation reversed the sign of a single term in equations (1–4) above (Table 1). Three of the eight simulations resulted in the agent preferring the bad action, four resulted in the agent preferring the good action, and one resulted in the agent selecting between them equally (because the action values alternated in sign on each update).

Both reversing the sign of r in equation (1a) or the sign of α in equation (3b) involve the δ term, which is the term interpreted as corresponding to DA firing in animals [10,12]. Ultimately these cases suggest that a reversed effect of DA on learned values of actions will cause an agent to consistently select the worst action.

However in these cases the agent becomes incapable of selecting any good action when a bad action is available. Since LNS patients can respond to positive reinforcement, it is clear that a sign reversal does not explain the data. Instead, taking the absolute value of the reward signal r , prediction error δ , or action values Q produces the effect of both the bad action and the good action having positive values. This may be interpreted as DA release or effects in the striatum increasing above baseline for both good and bad actions, thus reinforcing both.

This suggestion has support in the literature. The increased number of DA receptors and decreased number of DA transporters in LNS patients may make them hypersensitive to the effects of DA, including its putative role in reinforcement learning, and since DA levels in healthy animals increase following strong but not mild aversive stimuli[16], these stimuli may become rewarding in LNS patients. Also, agonists of the serotonin (5-HT) receptors found in striatum, 5-HT_{2A} and 5-HT_{2C}, increase local DA efflux[17]. In an animal model of LNS in which the DA system of neonatal rats is lesioned with 6-hydroxydopamine, rats show hyperinnervation of the striatum by 5-HT axons from the dorsal raphe nucleus[18]. In this animal model, 5-HT₂ receptors in striatal direct pathway neurons become coupled to the ERK1/2 MAP kinase pathway, which has a facilitatory effect on plasticity[19]. Further support for a general role of this system in self-injurious behavior comes from observations that risperidone, a strong DA and 5-HT_{2A} antagonist, often reduces self-injurious behavior in both autism[20] and mental retardation[21]. In sum, there are many potential mechanisms available by which these dysfunctions could produce strong dopaminergic or reinforcing effects in the striatum during aversive stimulation.

On a behavioral level, when a painful shock to the fingers had been used as a punishment for attempts at self-injury, the rate of self-injury attempts increased by two to three times in four of five subjects (with no change in rate for the fifth subject)[14]. The rate of attempts also increased when positive attention accompanied them[14], demonstrating that both positive and negative rewards paired with the attempts at self-injury caused an increase in responding.

We simulate common behavioral treatments of LNS[22] using equations 2–4, replacing equation 1 with

$$\delta = |r| + V(s') - V(s) \quad (1d)$$

where the absolute value of the reward represents the changes in DA and 5-HT input to the striatum in LNS.

Instead of attempting to assign many arbitrary values to the different behavioral treatments, we can consider a simpler, general case. In the healthy case where equation 1 is used, the action producing a negative reward is never taken (Fig. 1A). In contrast, equation 1d is used in the LNS model. For the case where the bad action produces a reward of -1 and the good action a reward of $+1$, there is a 50% chance of the model selecting either action (Fig. 1B, black line at the point $x = -1$). When the reward produced by the bad action is subsequently changed, the probability of selecting either action changes as well, decreasing as the reward approaches 0, increasing as the magnitude of the reward increases (Fig. 1B, black line). If the good action has reward of $+1.5$ instead, the overall probability of selecting the bad action is decreased for a wide range of associated rewards, though it has the same general shape (Fig. 1B, grey line).

Anderson and colleagues' baseline treatment used response prevention only[14]. This can be treated as a reward of a moderately small magnitude (Fig. 2a). When self-injury attempts are followed by a painful electrical shock, the reward is modeled as much larger and more negative, thus producing an increase in attempts at self-injury (Fig. 2b). Similarly, pairing positive attention with attempts at self-injury produces a positive reward of larger magnitude, which also increases the probability of self-injury (Fig. 2c).

Finally, as is consistently found in patients[14,22], when a desired behavior is specifically reinforced, producing a larger positive reward, the probability of selecting the bad action decreases (Fig. 2d).

Discussion

Essentially, since the hypothesized DA signal dysfunction treats any rewards as positive, differentially reinforcing any desired behavior over undesired behaviors is the only way to prevent the model from selecting undesired behaviors. In particular, the magnitude of the reward associated with an undesired behavior must be made as small as possible compared to the magnitude of the reward associated with desired behaviors. This offers an explanation for the maintenance of self-injurious behaviors and how they might initially be learned, but does not explain the cause of the initial behaviors themselves.

Interviews with the parents of two patients suggested that early self-injurious behavior appeared to be accidental, due to the poor motor control resulting from the disease, instead of purposeful [3]. From this we suggest that LNS patients are more likely to accidentally self-injure at early ages and that, because of dysfunctional DA input to their BG, the self-injury is reinforced instead of being stamped out. Similarly, it is plausible that the common forms of self-injury seen in adults could have initially occurred accidentally due to motor impairments [23].

One prediction of this model is that any arbitrary behavior should be reinforced through negative rewards (e.g. shocks or even verbal punishment) in LNS patients. This has been demonstrated for self-injurious behaviors [14] and the general occurrence of contrary behaviors in LNS patients has been reported [15], but it does not appear to have been specifically tested experimentally.

Modeling work has suggested that the error-related negativity (ERN) apparent in EEG activity arising from anterior cingulate cortex (ACC) is produced by midbrain DA signaling [24], which is supported by findings on a reduced ERN in patients with Parkinson's disease [25]. They proposed that reward prediction error-related decreased firing rate in DA neurons causes the ACC ERN. Combined with the present suggestion, this makes the prediction that LNS patients should not show an ERN in response to errors (or, further, to punishment or self-injurious behavior). Testing LNS patients for this prediction or any of the other predictions that follow from a dysfunction of BG function is complicated by the contrary or disobedient behavior of the patients [15].

Conclusion

We have shown that a model frequently used to interpret the functioning of the BG can be modified in a way consistent with the biology underlying LNS such that it is reinforced by normally aversive outcomes and that this model responds to previously tested behavioral treatments in the same way as patients. This supports the role of behavioral treatments in LNS patients and suggests a concrete explanation for the outcome of behavioral treatments, which may allow for the development of better treatments in the future.

Acknowledgements

This work was supported by NIMH MH60013, Silvio O. Conte Center grant NIMH MH71702, NSF SLC SBE 0354378 and NIDA DA16454 (part of the CRCNS program).

References

1. Lesch M, Nyhan WL. A Familial Disorder of Uric Acid Metabolism and Central Nervous System Function. *The American journal of medicine* 1964;36:561–570. [PubMed: 14142409]
2. Jinnah HA, Visser JE, Harris JC, Verdu A, Larovere L, Ceballos-Picot I, et al. Delineation of the motor disorder of Lesch-Nyhan disease. *Brain* 2006;129:1201–1217. [PubMed: 16549399]
3. Hall S, Oliver C, Murphy G. Self-injurious behaviour in young children with Lesch-Nyhan syndrome. *Developmental medicine and child neurology* 2001;43:745–749. [PubMed: 11730148]
4. Saito Y, Takashima S. Neurotransmitter changes in the pathophysiology of Lesch-Nyhan syndrome. *Brain & development* 2000;22(Suppl 1):S122–131. [PubMed: 10984673]
5. Nyhan WL. Dopamine function in Lesch-Nyhan disease. *Environmental health perspectives* 2000;108 (Suppl 3):409–411. [PubMed: 10852837]
6. Lloyd KG, Hornykiewicz O, Davidson L, Shannak K, Farley I, Goldstein M, et al. Biochemical evidence of dysfunction of brain neurotransmitters in the Lesch-Nyhan syndrome. *The New England journal of medicine* 1981;305:1106–1111. [PubMed: 6117011]
7. Ernst M, Zametkin AJ, Matochik JA, Pascualvaca D, Jons PH, Hardy K, et al. Presynaptic dopaminergic deficits in Lesch-Nyhan disease. *The New England journal of medicine* 1996;334:1568–1572. [PubMed: 8628337]
8. Saito Y, Ito M, Hanaoka S, Ohama E, Akaboshi S, Takashima S. Dopamine receptor upregulation in Lesch-Nyhan syndrome: a postmortem study. *Neuropediatrics* 1999;30:66–71. [PubMed: 10401687]
9. Wong DF, Harris JC, Naidu S, Yokoi F, Marengo S, Dannals RF, et al. Dopamine transporters are markedly reduced in Lesch-Nyhan disease in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93:5539–5543. [PubMed: 8643611]
10. Montague PR, Hyman SE, Cohen JD. Computational roles for dopamine in behavioural control. *Nature* 2004;431:760–767. [PubMed: 15483596]
11. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 1996;16:1936–1947. [PubMed: 8774460]
12. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science (New York, NY)* 1997;275:1593–1599.
13. Schultz W. Predictive reward signal of dopamine neurons. *Journal of neurophysiology* 1998;80:1–27. [PubMed: 9658025]
14. Anderson L, Dancis J, Alpert M. Behavioral contingencies and self-mutilation in Lesch-Nyhan disease. *Journal of consulting and clinical psychology* 1978;46:529–536. [PubMed: 670495]
15. Visser JE, Bar PR, Jinnah HA. Lesch-Nyhan disease and the basal ganglia. *Brain research* 2000;32:449–475.
16. Horvitz JC. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 2000;96:651–656. [PubMed: 10727783]
17. Lucas G, Spampinato U. Role of striatal serotonin2A and serotonin2C receptor subtypes in the control of in vivo dopamine outflow in the rat striatum. *Journal of neurochemistry* 2000;74:693–701. [PubMed: 10646521]
18. Brown P, Gerfen CR. Plasticity within striatal direct pathway neurons after neonatal dopamine depletion is mediated through a novel functional coupling of serotonin 5-HT2 receptors to the ERK 1/2 map kinase pathway. *The Journal of comparative neurology* 2006;498:415–430. [PubMed: 16871540]
19. Thomas GM, Huganir RL. MAPK cascade signalling and synaptic plasticity. *Nature reviews* 2004;5:173–183.
20. Gagliano A, Germano E, Pustorino G, Impallomeni C, D'Arrigo C, Calamoneri F, et al. Risperidone treatment of children with autistic disorder: effectiveness, tolerability, and pharmacokinetic implications. *Journal of child and adolescent psychopharmacology* 2004;14:39–47. [PubMed: 15142390]

21. Cohen SA, Ihrig K, Lott RS, Kerrick JM. Risperidone for aggression and self-injurious behavior in adults with mental retardation. *Journal of autism and developmental disorders* 1998;28:229–233. [PubMed: 9656134]
22. Olson L, Houlihan D. A review of behavioral treatments used for Lesch-Nyhan syndrome. *Behavior modification* 2000;24:202–222. [PubMed: 10804680]
23. Robey KL, Reck JF, Giacomini KD, Barabas G, Eddey GE. Modes and patterns of self-mutilation in persons with Lesch-Nyhan disease. *Developmental medicine and child neurology* 2003;45:167–171. [PubMed: 12613772]
24. Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review* 2002;109:679–709. [PubMed: 12374324]
25. Falkenstein M, Hielscher H, Dziobek I, Schwarzenau P, Hoormann J, Sunderman B, et al. Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport* 2001;12:157–161. [PubMed: 11201078]

Figure 1.

Probability of selecting an action over one with a fixed reward. The y-axis shows the probability that the "bad" action (which produces a reward specified by the x-axis) will be selected. The alternative action produces a fixed reward of +1 (black line) or +1.5 (grey line). A, in the normal model, an action that produces a negative reward is never taken. B, in the Lesch-Nyhan model, only the magnitude and not the sign of the reward produced by the action is important.

Figure 2.

Effects of behavioral treatments. The y-axis shows the probability that the “bad” action (which produces a reward specified by the x-axis) will be selected. The alternative action produces a fixed reward of +1 (black line) or +1.5 (grey line). From a baseline probability of taking the bad action (a), making the reward more negative (punishment; (b)) or making the reward positive and of a larger magnitude (positive attention; (c)) both increase the probability of selecting the bad action. By increasing the reward of the good action to +1.5 (differential reinforcement), the probability of taking the bad action decreases (d).

Table 1

Each row represents a set of simulations using equations (1–4) but with the modified equation from the table replacing the corresponding original equation (e.g. the first row represents a simulation using equations 1a, 2, 3, and 4). The “bad” action produced a reward of -1 and the “good” action produced a reward of $+1$.

Probability of selecting “bad” action as simulation equations vary.	
Modified Equation	Approximate Probability of Selecting “Bad” Action
(1a) $\delta = (-1)r + \gamma V(s') - V(s)$	0.96
(1b) $\delta = r + (-1)\gamma V(s') - V(s)$	0.04
(1c) $\delta = r + \gamma V(s') - (-1)V(s)$	0.02
(2a) $V(s) \leftarrow (-1)V(s) + \alpha\delta$	0.11
(2b) $V(s) \leftarrow V(s) + (-1)\alpha\delta$	0.02
(3a) $Q(s, a) \leftarrow (-1)Q(s, a) + \alpha\delta$	0.50
(3b) $Q(s, a) \leftarrow Q(s, a) + (-1)\alpha\delta$	0.96
(4a) $\Pr(a s) = \frac{\exp(-Q(s, a))}{\sum_{a'} \exp(-Q(s, a'))}$	0.96