8. Lee, H.W. *et al. J. Neurosci.* **28**, 14546–14556 (2008).
9. Collins, M.O. & Grant, S.G. *Subcell. Biochem.* **43**, 185–207 (2007).
10. Tu, J.C. *et al. Neuron* **23**, 583–592 (1999).
11. Kitano, J. *et al. J. Neurosci.* **22**, 1280–1289 (2002).
12. Paquet, M. *et al. J. Biol. Chem.* **281**, 29949–29961 (2006).
13. Adwanikar, H. *et al. Pain* **111**, 125–135 (2004).
14. Tappe-Theodor, A. *et al. Mol. Pain* **7**, 38 (2011).
15. Zhuang, Z.Y. *et al. J. Neurosci.* **24**, 8300–8309 (2004).

Reward and punishment illuminated

Joseph J Paton & Kenway Louie

How do outcomes affect future behavior? A study using precise optogenetic stimulation finds that learning from positive reinforcement is mediated by striatal pathways distinct from those that mediate learning from punishment.

© 2012 Nature America, Inc. All rights reserved.

npg

When a new cafe opens, how do we decide to stop for a cup of coffee? Absent any pre-existing knowledge, we have to taste a coffee or two and use that information to guide our morning routine. One hallmark of adaptive behavior is the ability to learn from the outcomes of actions, whether those results are positive or negative. This importance of outcomes in shaping behavior is codified in Thorndike's law of effect: behaviors associated with satisfaction or discomfort will be more or less likely to recur, respectively. Decades of lesion, electrophysiology and functional imaging studies have outlined a broad network of brain areas involved in learning about reward and punishment, but how this processing is integrated with action selection remains unknown. In this issue of *Nature Neuroscience*, Kravitz *et al.*[1] report that two distinct pathways in the basal ganglia, a subcortical system that is essential for motor control, differentially mediate reward and punishment.

A critical tool for examining the neural basis of reinforcement has been the targeted activation of specific brain areas. Electrical stimulation has long been known to elicit behavioral phenomena, ranging from simple percepts and movements to emotions and vivid memories, depending on the targeted brain region[2,3]. In learning protocols, stimulation itself can serve as reinforcement, driving animals to seek or avoid further stimulation[4]. A wealth of brain stimulation reward studies have identified a network of subcortical areas involved in reward processing, notably the lateral hypothalamus, medial forebrain bundle and mesolimbic dopaminergic system[5]. The firing of dopaminergic neurons may be critical for learning, signaling a reward prediction error representing whether the moment at hand is better or worse than expected[6]. An intuition behind this kind of teaching signal is that if predictions turn out to be inaccurate in a given situation, this is a useful sign that the system should update its valuations to improve future predictions.

One substantial drawback of electrical stimulation, however, is nonspecific activation in the vicinity of the stimulating electrode, including all local neuronal cell bodies and neuronal axons passing by en route to distant locations. Such nonspecificity is particularly problematic for studying areas without a gross functional architecture; for example, nuclei where neurons projecting to different targets or encoding different information are closely intermingled. The recent arrival of optogenetic techniques, which combine light-activated spiking activity with genetic localization to specific cell types, represents a substantial advance in the specificity of targeted activation. Kravitz *et al.*[1] used optogenetic control to selectively target two distinct populations of neurons in the striatum, the major input nucleus of the basal ganglia.

Roughly 95% of the neurons in the striatum are GABAergic projection cells called medium spiny neurons (MSNs), so named for their medium size and the dense distribution of spines along their dendrites. MSNs receive excitatory input from layer 5 pyramidal neurons in almost all areas of cortex, as well as from the thalamus, and modulatory input from dopaminergic neurons in the substantia nigra pars compacta and/or the ventral tegmental area in the midbrain. Striatal MSNs can be subdivided into two subclasses on the basis of projection target and expression of dopamine receptor types. Direct pathway striatonigral MSNs (dMSNs) express D1 dopamine receptors and project directly to basal ganglia output nuclei: the internal segment of the globus pallidus (GPi) and/or the substantia nigra pars reticulata (SNr). Indirect pathway striatopallidal MSNs (iMSNs) express D2 dopamine receptors and represent the first stage of a more indirect route to basal ganglia output, terminating primarily in the external segment of the globus pallidus (GPe).

The direct and indirect pathways comprise the two fundamental opposing forces in the classic model of basal ganglia motor control[7,8]. Excitatory corticostriatal input to direct pathway dMSNs increases striatonigral inhibition of GPi and SNr activity, disinhibiting thalamocortical projections and facilitating movement (**Fig. 1a**). In the indirect pathway, striatal activity in iMSNs acting through the GPe and subthalamic nucleus results in a net inhibition of thalamocortical activity and suppression of movement (**Fig. 1b**). Action selection is thought to be implemented by a striatal competition between actions specified by corticostriatal inputs, mediated by balanced direct and indirect pathway activity. Although recent anatomical and functional evidence suggests that this scheme may be oversimplified, major motor pathologies can be explained by an imbalance in striatal information processing. In Parkinson's disease, loss of dopaminergic input leads to an overactivity of indirect versus direct pathway activity, resulting in a poverty of movement; in Huntington's disease, loss of indirect pathway activity removes inhibitory control, resulting in an excess of abnormal movements.

In addition to action selection, growing evidence also implicates the basal ganglia in reward learning[9,10]. Theoretical models suggest that learning from positive and negative outcomes could be functionally segregated via the direct and indirect pathways, suggesting that reward and punishment may be mediated by separate anatomic systems. Given its reward-related signaling and dense innervation of the striatum, dopaminergic input is likely to be critical. Indeed, D1 type and D2 type receptors in the striatum have opposite effects on cell excitability: activation of D1 receptors increases excitability of dMSNs and activation of D2 receptors decreases excitability of iMSNs, both acting via G protein–coupled changes in responsiveness to glutamatergic input. The phasic release of dopamine around better-than-expected behavioral events would be expected

Joseph J. Paton is at the Champalimaud Neuroscience Programme, Lisbon, Portugal. Kenway Louie is at the Center for Neural Science, New York University, New York, New York, USA. e-mail: klouie@cns.nyu.edu

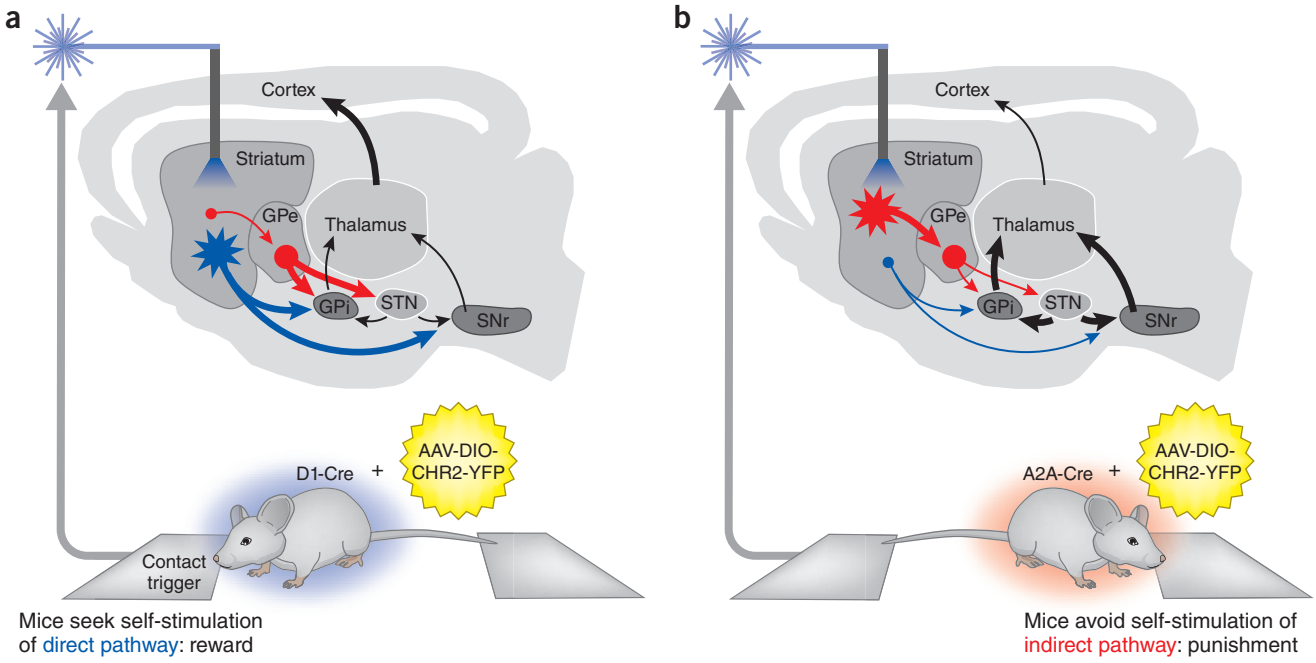

Figure 1 Simplified schematic of basal ganglia anatomy and circuit response to optogenetic stimulation. (**a**,**b**) Kravitz *et al.*[1] expressed the light-sensitive cation channel Channelrhodopsin2 (ChR2) in dMSNs (**a**) and iMSNs (**b**) by injecting adeno-associated virus (AAV) containing a *loxP*-flanked inverted *ChR2* construct into the dorsomedial striatum of transgenic mice expressing Cre recombinase in dMSNs (D1-Cre) or iMSNs (A2A-Cre), respectively. They then placed the mice in an environment with two contact devices. One contact triggered a brief pulse of laser light delivered to the striatum through an optical fiber. The other contact was inactive. Activation of MSNs is thought to change the activity at many points in the downstream circuitry, ultimately leading to a decrease or increase in inhibitory output from the basal ganglia when stimulating dMSNs (**a**) or iMSNs (**b**), respectively. Arrow thickness indicates predicted relative activity in different projection pathways in response to the two stimulation conditions. Blue, direct pathway; red, indirect pathway. STN, subthalamic nucleus. Areas outlined in black send inhibitory projections. Those outlined in white send excitatory projections.

to increase the effectiveness of inputs to direct pathway MSNs and to decrease the effectiveness of inputs to indirect pathway MSNs; phasic decreases of dopamine around worse-than-expected events might do the opposite. In their study, Kravitz *et al.*[1] address a specific and intriguing hypothesis: direct pathway activation mimics reward and indirect pathway stimulation mimics punishment.

The authors used a viral delivery strategy to target the light-activated protein Channelrhodopsin2 to either dMSNs or iMSNs in the dorsomedial striatum of transgenic mice, allowing selective activation of either the direct or indirect pathway. Analogously to classic brain stimulation reward studies, the authors provided laser-mediated activation of the targeted neurons in an operant manner, whenever the animal touched a sensitive capacitive trigger. Mice given direct pathway stimulation showed a significant bias toward the laser-paired trigger in comparison with a control trigger unpaired with stimulation. In contrast, mice given indirect pathway stimulation exhibited an opposite bias, avoiding the laser-paired trigger. Together, these findings suggest that striatal activation is sufficient for mediating both reinforcement and punishment, depending on the subpopulation of stimulated neurons.

A critical question is how these reinforcement effects are related to the well-documented role of the basal ganglia in motor control. Previously, in an experiment using the same optogenetic techniques and mouse lines, direct pathway stimulation increased locomotion and decreased freezing, whereas indirect pathway stimulation produced a parkinsonian state, with decreased locomotion, increased freezing and bradykinesia[11]. Such broad stimulation-induced movement changes raise the possibility that the behaviors observed in the current study arise from simple changes in overall motor activity. However, Kravitz *et al.*[1] found that striatal activation induced few changes in movement parameters, possibly owing to short laser stimulation durations. Furthermore, although trigger biases emerged gradually over the course of initial training sessions, they were immediately observed in subsequent sessions, consistent with a learned behavior rather than a motor confound.

Which neural systems mediate this simulation-induced reinforcement learning? In the standard model of basal ganglia function, the rich dopaminergic innervation of the striatum can influence action selection in two primary ways[12,13]. First, dopamine indirectly excites dMSNs and inhibits iMSNs, primarily by altering the responsiveness to glutamatergic input. Second, phasic dopamine bursts may induce long-lasting changes in striatal plasticity, including long-term potentiation via D1 receptor activation and long-term depression via D2 receptor activation. Notably, by targeting dopamine receptor–expressing striatal MSNs, Kravitz *et al.*[1] effectively bypassed dopaminergic signaling. To confirm that their reinforcement effects were independent of dopamine, the authors repeated the operant conditioning experiments after injecting a combination of D1 and D2 receptor antagonists. Although these antagonists reduced overall locomotor activity, the positive and negative trigger biases associated with direct and indirect pathway stimulation remained intact.

These findings suggest that the activation of striatal pathways can mediate the effects of reinforcement independent of dopaminergic signaling, but the site of learning remains an open question. Although learning appears to be independent of dopamine-induced plasticity at synapses on MSNs, co-occurrence of striatal input and stimulation might change efficacy at corticostriatal and thalamostriatal synapses through other, dopamine-independent mechanisms, such as glutamate

© 2012 Nature America, Inc. All rights reserved.

receptor–driven plasticity[14,15]. Alternatively, optogenetic activation may induce plasticity downstream of striatal activation, either in the basal ganglia at subthalamic nucleus, GPe or GPi synapses, or possibly further along in thalamic or cortical areas. Notably, Kravitz *et al.*[1] found an asymmetry in the temporal longevity of stimulation effects, with positive reinforcement effects outlasting the transient punishment effects of activation; such differences provide an intriguing starting point for future studies into the neural locus of reinforcement learning. In a final experiment, the authors show that direct and indirect pathway stimulation also changes behavior in a place preference task, suggesting that striatal activation may provide a reinforcement signal that generalizes to context as well as action.

Overall, the results of Kravitz *et al.*[1] highlight a fundamental point about decision-making: selecting an action is never truly independent of reward learning. Functionally, the learned values of different options is a crucial element of the action selection process. Neurally, action selection and reinforcement learning appear to be implemented in the same striatal circuitry, with distinct functional compartments processing rewarding versus aversive outcomes. Understanding the exact nature and mechanism of this relationship between reinforcement and action will be a critical avenue for further research.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Kravitz, A.K., Tye, L.D. & Kreitzer, A.C. *Nat. Neurosci.* **15**, 816–818 (2012).
2. Taylor, C.S. & Gross, C.G. *Neuroscientist* **9**, 332–342 (2003).
3. Penfield, W. *The Excitable Cortex in Conscious Man* (C.C. Thomas, 1958).
4. Olds, J. & Milner, P. *J. Comp. Physiol. Psychol.* **47**, 419–427 (1954).
5. Wise, R.A. *Neuron* **36**, 229–240 (2002).
6. Schultz, W., Dayan, P. & Montague, P.R. *Science* **275**, 1593–1599 (1997).
7. Albin, R.L., Young, A.B. & Penney, J.B. *Trends Neurosci.* **12**, 366–375 (1989).
8. Graybiel, A.M. *Curr. Biol.* **10**, R509–R511 (2000).
9. Wickens, J.R., Reynolds, J.N. & Hyland, B.I. *Curr. Opin. Neurobiol.* **13**, 685–690 (2003).
10. Frank, M.J. *Curr. Opin. Neurobiol.* **21**, 381–386 (2011).
11. Kravitz, A.V. *et al. Nature* **466**, 622–626 (2010).
12. Surmeier, D.J., Ding, J., Day, M., Wang, Z. & Shen, W. *Trends Neurosci.* **30**, 228–235 (2007).
13. Kreitzer, A.C. & Malenka, R.C. *Neuron* **60**, 543–554 (2008).
14. Dang, M.T. *et al. Proc. Natl. Acad. Sci. USA* **103**, 15254–15259 (2006).
15. Koralek, A.C., Jin, X., Long, J.D. II, Costa, R.M. & Carmena, J.M. *Nature* **483**, 331–335 (2012).

© 2012 Nature America, Inc. All rights reserved.

Squaring cortex with color

Brian A Wandell & E J Chichilnisky

A long-standing puzzle has been the seeming inconsistency between neuronal responses in primary visual cortex to colored stimuli and the elementary perceptual attributes of color vision. Nonlinear analysis resolves this paradox.

Color perception begins with light absorption by three types of cone photoreceptors in the retina. Cone signals are processed by visual system circuitry to produce the familiar perceptual attributes of color appearance. In the past several decades, studies of this circuitry have revealed consistency between perceptual experiments and the responses of neurons in the retina and thalamus. However, responses of neurons in primary visual cortex (area V1) to colored stimuli have been difficult to reconcile with these findings. Now, work by Horwitz and Hass[1] that accounts for certain nonlinear processing characteristics of V1 neurons points the way to a unified understanding.

A fundamental aspect of color vision is that an appearance match can be arranged between any target light and an appropriate combination of three fixed primaries, even though the spectra of the target and combined primaries differ greatly. The detailed and quantitative characterization of this phenomenon, known as color matching, was a great achievement of twentieth century science. Our understanding of color matching is the basis for all technologies relating to color imaging[2]. For example, modern color displays repeat the color matching experiment 60 times a second in millions of pixels to provide realistic scene renditions. Quantitative measurements of the spectral sensitivity of the three types of cones secured the connection between the biology of light transduction and color matching[3].

But color matching does not explain color appearance. We can match the appearance of two lights without saying what either looks like, just as we can match the weight of two objects without knowing what either is made of. Consequently, modern color science has focused on determining how color appearance is derived from the signals in the three types of cones. The first major step was the remarkable discovery of a simple empirical rule of color appearance: certain colors occur in combination, whereas others do not. For example, the color orange appears both reddish and yellowish. But there are no colors that simultaneously appear red and green or that simultaneously appear blue and yellow. These forbidden color pairs were discovered by Hering[4] and developed into the theory of color opponency by Hurvich and Jameson[5] and others[6,7].

What is the neural basis of color opponency? Svaetichin[8] and DeValois[9] discovered that neurons in the retina and the lateral geniculate nucleus (LGN), the subcortical relay station en route to V1, encode color signals as sums and differences of cone signals. For example, certain retinal and LGN neurons are excited by inputs from the cones that are sensitive to long wavelengths and are suppressed by inputs from the cones that are sensitive to middle wavelengths. If increases in the activity of these cells encode red and decreases encode green, such neurons could report red or green, but not both, which is consistent with the observed exclusivity of red and green percepts. A similar situation holds for blue-yellow opponent circuitry. Neural responses in the LGN were grouped into three categories that roughly matched the three categories of color-opponent theory[10]: red-green, blue-yellow and light-dark. Thus, to a first approximation, circuitry in the subcortical visual system implements the perceptual phenomenon of color opponency in physiology and behavior. The agreement between physiology and behavior falls short of the precision of color matching, but it is not so far off as to cause any serious alarm.

Cortical neurons, however, have refused to join the party. Over the past several decades, studies of color processing in visual cortex have produced confusing and seemingly contradictory results. The earliest studies in V1 suggested that there were very few color-opponent cells[11]. Subsequent studies found

Brian A. Wandell is in the Department of Psychology, Stanford University, Stanford, California, USA, and E.J. Chichilnisky is at The Salk Institute for Biological Studies, La Jolla, California, USA.
e-mail: wandell@stanford.edu