NHE6 itself is trafficked through the endosomal recycling pathway. Rack1 interacts with the C terminus of NHE6, NHE7, and NHE9. Knockdown of Rack1 caused a decrease in cell surface levels of NHE6, while an increase in endosomal-associated NHE6 led to elevated endosomal pH, indicating the importance of maintaining NHE6 levels between intracellular compartments and the plasma membrane. These data suggest that a tightly regulated distribution of NHE6 is required for proper polarized membrane trafficking and maintenance of cell polarity (Ohgaki et al., 2011). Similar trafficking control is surely operative in neurons to ensure the correct distribution of proton pumps and exchanger and thereby the correct acidification of the endosome. Much still needs to be learned about the molecular mechanisms that achieve this and what the functional consequences are of loss of proper organelle acidification. Exciting progress is now linking disturbances in acidification to neurodegeneration (Li and DiFiglia, 2012; Nixon et al., 2001; Wang and Hiesinger, 2012) and, in this most recent paper, to autism-related neurodevelopmental disorders.

REFERENCES

Ascaño, M., Bodmer, D., and Kuruvilla, R. (2012). Trends Cell Biol. *22*, 266–273.

Deane, E.C., Ilie, A.E., Sizdahkhani, S., Das Gupta, M., Orlowski, J., and McKinney, R.A. (2013). J. Neurosci. *33*, 595–610.

Gilfillan, G.D., Selmer, K.K., Roxrud, I., Smith, R., Kyllerman, M., Eiklid, K., Kroken, M., Mattingsdal, M., Egeland, T., Stenmark, H., et al. (2008). Am. J. Hum. Genet. *82*, 1003–1010.

Huotari, J., and Helenius, A. (2011). EMBO J. *30*, 3481–3500.

Jan, Y.-N., and Jan, L.Y. (2010). Nat. Rev. Neurosci. *11*, 316–328.

Li, X., and DiFiglia, M. (2012). Prog. Neurobiol. *97*, 127–141.

Nixon, R.A., Mathews, P.M., and Cataldo, A.M. (2001). J. Alzheimers Dis. *3*, 97–107.

Ohgaki, R., van IJzendoorn, S.C.D., Matsushita, M., Hoekstra, D., and Kanazawa, H. (2011). Biochemistry *50*, 443–450.

Ouyang, Q., Lizarraga, S.B., Schmidt, M., Yan, U., Gong, J., Elisor, D., Kauer, J.A., and Morrow, E.M. (2013). Neuron *80*, this issue, 97–112.

Schwede, M., Garbett, K., Mirnics, K., Geschwind, D.H., and Morrow, E.M. (2013). Mol. Psychiatry. Published online March 19, 2013. http://dx.doi.org/10.1038/mp.2013.28.

Scott, C.C., and Gruenberg, J. (2011). Bioessays *33*, 103–110.

Strømme, P., Dobrenis, K., Sillitoe, R.V., Gulinello, M., Ali, N.F., Davidson, C., Micsenyi, M.C., Stephney, G., Ellevog, L., Klungland, A., and Walkley, S.U. (2011). Brain *134*, 3369–3383.

Walsh, C.A., Morrow, E.M., and Rubenstein, J.L.R. (2008). Cell *135*, 396–400.

Wang, D., and Hiesinger, P.R. (2012). Autophagy *8*, 711–713.

Yap, C.C., and Winckler, B. (2012). Neuron *74*, 440–451.

Exploiting Exploration: Past Outcomes and Future Actions

Kenway Louie[1,*]
[1]Center for Neural Science, New York University, 4 Washington Place, Room 809, New York, NY 10003, USA
*Correspondence: klouie@cns.nyu.edu
http://dx.doi.org/10.1016/j.neuron.2013.09.016

Applying past knowledge to future actions is crucial for adaptive choice behavior. Here, in this issue of *Neuron*, Donahue et al. (2013) show that reward enhances neural coding reliability for actions in a network of frontal and parietal brain areas.

Consider a soccer player lining up for a penalty kick, who knows from past experience that the goalie has a slight bias for rightward saves but only at the end of a match. To use that information, he must weigh the context, appropriately value different alternatives, and select and execute an action. Thus, the process of applying prior knowledge to future behavior involves a number of related cognitive functions—including valuation, memory, action selection, and cognitive control—necessary for adaptive decision making in a dynamic environment.

An influential framework integrating these processes arises from computational theories of machine learning (Sutton and Barto, 1998). The core idea in these reinforcement learning (RL) models is that agents acquire information about the value of actions through interaction with the environment, using reward to guide the learning process. To update the value of actions, such models employ error-driven learning using a quantity known as reward prediction error (RPE), the difference between reward received and reward expected. For example, actions that produce reward that is better than expected have their associated values increased. In stable environments, this procedure produces value estimates that converge appropriately to the average reward. Neuroscientific interest in RL emerged with the discovery that midbrain dopamine neuron activity in classical and operant conditioning tasks carries an

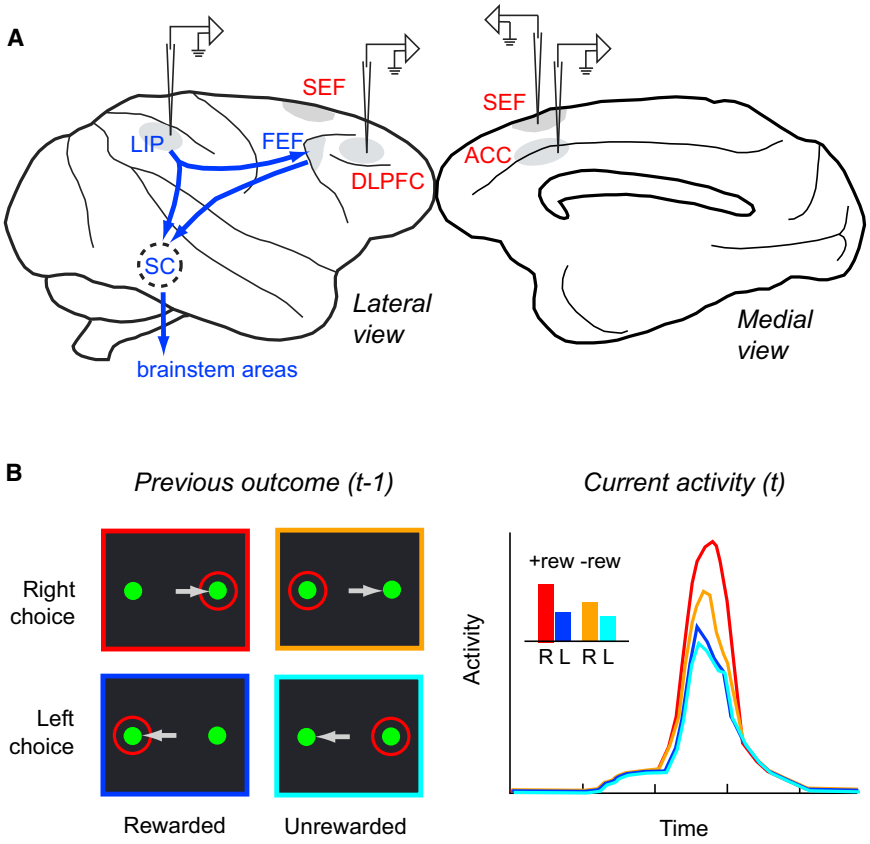

Figure 1. Neural Activity Underlying Strategic Exploration

(A) Brain regions involved in saccadic decision making. Donahue et al. (2013) report analyses of neural data from multiple cortical regions, including those involved in saccade selection and execution (LIP) and those involved in postaction processing and cognitive control (ACC, SEF, and DLPFC).

(B) The effect of past actions and outcomes on neural activity during strategic exploration. Left: representation of possible outcomes in the matching pennies game (red circle, computer choice; arrow, monkey choice). Right: schematic depicting interactive effect of previous reward and action on current neural activity. This interaction produces an enhanced neural discrimination of past actions when reward was previously received (inset).

RPE signal (Schultz et al., 1997). Subsequent work has successfully applied the RL framework in a variety of brain systems and behavioral paradigms to characterize value- and choice-related neural activity and value-guided decision behavior (Lee et al., 2012).

Reinforcement learning works remarkably well in stable environments but faces an additional important challenge when the state of the world is uncertain and changing (Sutton and Barto, 1998). Given an evaluation of the possible actions, the goal of an agent is to exploit current knowledge in choosing the highest valued option. However, changing conditions in a dynamic environment necessitate an exploration of nonoptimal alternatives in order to maintain accurately updated values. Solving this tradeoff between exploration and exploitation is a fundamental problem in learning through reinforcement. A particular question of interest is how the brain switches from exploitative behavior, which is a natural byproduct of a value-guided decision system, to strategic exploratory behavior, which forgoes current value maximization for a more global optimality. Recent progress has identified neural substrates involved in exploration, such as the neuromodulatory noradrenergic system (Usher et al., 1999) and frontopolar cortex (Daw et al., 2006), but the full extent of cortical circuits involved in strategic exploration is unknown.

In this issue of *Neuron*, Donahue et al. (2013) examine the relationship between strategic exploration and a network of cortical regions related to saccade selection, execution, and postsaccade processing. Action selection in the eye movement system has long been a model of neurobiological decision making (Glimcher, 2003), and lesion and electrophysiology studies have identified core sensorimotor structures involved in the decision process (Figure 1A). A key structure in this network is the lateral intraparietal (LIP) area, which receives afferents from higher-order sensory areas and displays both sensory and motor modulation. Consistent with a central role in decision making, saccade-selective activity in LIP represents the information necessary for decision formation, for example, accumulating evidence for a given response in perceptual discrimination tasks (Shadlen and Newsome, 2001).

LIP efferents project to the frontal eye field (FEF) and superior colliculus (SC), and together this network (along with the caudate in the basal ganglia) plays an essential role in saccade selection and execution. FEF and SC are necessary for saccade generation: saccades are initiated only when movement-related activity reaches a fixed threshold, microstimulation in these structures elicits fixed vector saccades, and lesions disrupt saccade initiation. Consistent with the anatomy, LIP appears to play a more upstream role. While lesions in LIP leave saccades to single targets relatively intact, they produce substantial deficits in target selection from multiple alternatives. Importantly, action value information strongly modulates neural activity in these areas during the choice process, consistent with an integrated evaluation and decision-making network (Glimcher, 2003). Note that this system, which selects saccades based on value, is designed to implement exploitation behavior.

However, there are a number of additional brain areas anatomically and functionally linked to these core sensorimotor circuits that play a different, less transparent role in choice behavior. These areas include three interconnected regions (among others) in frontal cortex: the supplementary eye field (SEF), anterior cingulate cortex (ACC), and dorsolateral prefrontal cortex (DLPFC). These

areas are related to saccades but not necessary for action initiation or execution; instead, activity in these regions represent a variety of error- and reward-related signals that may be involved in performance monitoring and executive control of the gaze process (Ito et al., 2003; Schall et al., 2002; Stuphorn et al., 2010). Neural activity in these areas, most notably in SEF and ACC, often occurs after action completion, consistent with a role in reward processing and outcome-based updating. Similar evaluative signals occur in humans, where strong negative potentials are recorded over medial frontal cortex when errors are made in simple behavioral tasks (Gehring et al., 1993). In contrast to the core oculomotor network defined by LIP and FEF, postaction processing in these frontal areas suggests a role in executive control and a potential involvement in regulating exploratory behavior.

In this study, Donahue et al. (2013) examine the neural basis of strategic exploration by taking advantage of an impressive data set of recordings from SEF, DLPFC, ACC, and LIP neurons and using two behavioral tasks designed to elicit either exploitation or exploration. To elicit exploitation behavior, they used a simple visual search task, where the location of the rewarded target was explicitly cued in each trial. In this task, reward was determined by a fixed rule and monkeys simply had to choose the high-value, cued target. To elicit exploration behavior, they used a competitive game known as matching pennies. In this task, played against a computer opponent, monkeys chose between two identical targets. Much like the soccer player taking a penalty kick, reward outcome depended on the behavior of the opponent: the monkey was rewarded only if he chose the same target chosen by the computer (revealed after the animal's choice).

Importantly, the computer opponent employed an algorithm that took advantage of any statistical biases evident in the animal's behavior. Thus, to achieve optimal reinforcement rates, the monkey should on average choose each target equally and with independent probability, irrespective of past choices and outcomes. Using this form of competitive game has two experimental advantages. First, by penalizing behavioral biases, this task encourages strategic exploration rather than deterministic behavior. Second, the resulting stochastic behavior dissociates past actions and reward from future choices, enabling the experimenters to determine whether neural activity reflects the influence of previous knowledge or current action planning.

Donahue et al. (2013) find that previous reward and actions influence activity during the matching pennies task in all four cortical regions examined but with some notable and important differences between areas. In a given trial, during the time before the monkey made a choice, a significant fraction of neurons in all four areas signaled the choice and reward outcome in the previous trial. Notably, neurons in SEF, DLPFC, and LIP—but not ACC—also coded the interaction between previous reward and choice (Figure 1B). This interaction reflects a gating of action coding by reward, such that neural discrimination between past actions is enhanced if reward was received (inset). Because past and future choices were dissociated in this task, Donahue et al. (2013) further show that this enhanced discriminability reflects information about past but not upcoming choice.

Intriguingly, Donahue et al. (2013) find that the SEF may play a particularly important role in governing exploratory behavior. While performance in the matching pennies task approached optimal randomness, the monkeys showed a slight but significant bias in their behavior. Specifically, they adopted an asymmetric win-stay lose-switch strategy, repeating previous choices if rewarded and switching targets if unrewarded in the previous trial. Their slight tendency to win-stay more often than lose-shift produced a small and fluctuating bias to choose the same target in successive trials. Notably, Donahue et al. (2013) found that switching behavior was significantly correlated with the reward-driven improvements in neural action decoding but only for the SEF. Furthermore, enhanced discriminability in SEF was largely attenuated during the visual search task, suggesting that it may play a unique role in guiding exploration behavior.

These findings raise important questions about how these different cortical regions interact in reinforcement-guided behavior. Enhanced SEF action coding following reward appears to facilitate subsequent exploratory switching behavior, but exactly how it does so is not known. One likely candidate is the strong projections SEF sends to the FEF. Microstimulation of SEF neurons can produce either excitatory or suppressive effects on FEF-mediated saccade initiation, consistent with a contextual form of executive control sensitive to task demands (Stuphorn and Schall, 2006). Thus, SEF may drive exploratory behavior by proactively influencing the saccade selection process in FEF, perhaps by overriding the default exploitation behavior driven by reinforcement learning. Characterizing the nature of this interaction will be an important focus of future research.

Ultimately, these results point toward a more nuanced view of reinforcement learning in the brain. Traditional RL algorithms, including many of those used to study decision-related neural activity, focus on learning the values of actions and choose according to previously received reward. In contrast to such model-free RL, increasing work has focused on model-based learning strategies, which carry an internal model of the world and attempt to learn the sequential contingencies of events, actions, and reward (Doll et al., 2012). In the complicated dynamics of a competitive game, reward is determined not by the choice of a particular action but by a sequence of actions. Thus, a monkey playing matching pennies must learn strategies rather than specific actions. This complexity may explain why circuits like ACC and DLPFC, which display significant choice- and reward-related activity related to value-guided behavior, apparently contribute little to strategic exploration behavior. Much work remains to be done in characterizing the interconnected brain regions responsible for exploitation, exploration, and their relative balance. These current findings provide an important roadmap for future research at the intersection of reinforcement learning and strategic behavior.

REFERENCES

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Nature *441*, 876–879.

Doll, B.B., Simon, D.A., and Daw, N.D. (2012). Curr. Opin. Neurobiol. *22*, 1075–1081.

Donahue, C.H., Seo, H., and Lee, D. (2013). Neuron *80*, this issue, 223–234.

Gehring, W.J., Gross, B., Coles, M.G.H., Meyer, D.E., and Donchin, E. (1993). Psychol. Sci. *4*, 385–390.

Glimcher, P.W. (2003). Annu. Rev. Neurosci. *26*, 133–179.

Ito, S., Stuphorn, V., Brown, J.W., and Schall, J.D. (2003). Science *302*, 120–122.

Lee, D., Seo, H., and Jung, M.W. (2012). Annu. Rev. Neurosci. *35*, 287–308.

Schall, J.D., Stuphorn, V., and Brown, J.W. (2002). Neuron *36*, 309–322.

Schultz, W., Dayan, P., and Montague, P.R. (1997). Science *275*, 1593–1599.

Shadlen, M.N., and Newsome, W.T. (2001). J. Neurophysiol. *86*, 1916–1936.

Stuphorn, V., and Schall, J.D. (2006). Nat. Neurosci. *9*, 925–931.

Stuphorn, V., Brown, J.W., and Schall, J.D. (2010). J. Neurophysiol. *103*, 801–816.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction. (Cambridge: MIT Press).

Usher, M., Cohen, J.D., Servan-Schreiber, D., Rajkowski, J., and Aston-Jones, G. (1999). Science *283*, 549–554.

The Sacred Disease: The Puzzling Genetics of Epileptic Disorders

Gaia Novarino,[1] Seung Tae Baek,[1] and Joseph G. Gleeson[1,*]
[1]Howard Hughes Medical Institute, Department of Neurosciences, University of California, San Diego, La Jolla, CA 92093, USA
*Correspondence: jogleeson@ucsd.edu
http://dx.doi.org/10.1016/j.neuron.2013.09.019

In the September 12, 2013 issue of *Nature*, the Epi4K Consortium (Allen et al., 2013) reported sequencing 264 patient trios with epileptic encephalopathies. The Consortium focused on genes exceptionally intolerant to sequence variations and found substantial interconnections with autism and intellectual disability gene networks.

This is a particularly exciting time for the genetics of the epilepsies, especially considering its sordid history. For centuries, epilepsy was considered a sign of a supernatural presence, often of a demonic variety. More than 2,000 years ago, however, Hippocrates warned that epilepsy was no more sacred than other diseases and that its origins may lie in heredity. Yet, for many years, the epilepsies were not considered of genetic origin and up until the 20th century epileptics were socially isolated. The strong genetic component of epilepsy was suggested by observation of familial aggregation, confirmed by several monozygotic-dizygotic twin studies, and in the 1990s, the first epilepsy-specific genes were cloned, encoding ion channels and neurotransmitter receptors. Dominant and recessive Mendelian inheritance patterns have been verified for several of these, but it became clear that these mutations are subject to partial penetrance (i.e., not every person with the mutation will develop epilepsy) and that even if epilepsy develops in a carrier, not everyone with the mutation will display the same form of epilepsy.

Other types of genetic alterations, such as de novo (sporadic) and somatic (limited to specific brain areas) mutations were long suspected to contribute to epilepsy but not validated until recently. Whole-exome sequencing (WES) provided a new tool to understand this multifactorial disorder, allowing a window into the genetic architecture that for the first time did not require pedigree and linkage analysis.

Over the course of several years and with generous support from the NINDS, the Epilepsy Genome Phenome Program (http://www.epgp.org) recruited hundreds of patients and families from an international network of 27 clinical centers in the U.S., Europe, and Australia. The goal of the EPGP program was to enroll 1,500 families in which two or more affecteds displayed epilepsy and 750 individuals with epileptic encephalopathies (EEs). EEs are a group of progressive partially overlapping neurological syndromes in which patients, usually young children, present with psychomotor dysfunctions and concurrent severe clinical epilepsy, often with infantile spasms, Lennox-Gastaut syndrome, polymicrogyria, or periventricular heterotopias. The EPGP cohorts were recruited, meticulously phenotyped, and subsequently underwent DNA exome sequencing through the Epi4K consortium, again funded by the NINDS. The results of the first sequencing effort of the EEs were recently published in the journal *Nature* (Allen et al., 2013).

Following the idea that clinical homogeneity corresponds to genetic homogeneity, Allen et al. (2013) focused on two well-described EEs: infantile spasms and Lennox-Gastaut syndrome. Collectively, they sequenced 264 patients and their