CellPress

Are You Smarter Than a Teenager? Maybe Not When It Comes to Reinforcement Learning

Brynne C. DiMenichi[1] and Elizabeth Tricomi[1,*]
[1]Department of Psychology, Rutgers University–Newark, 101 Warren Street, Newark, NJ 07102, USA
*Correspondence: etricomi@rutgers.edu
http://dx.doi.org/10.1016/j.neuron.2016.09.043

Adolescence is a time of tumultuous behavior that may result, in part, from brain circuitry that enhances reward seeking. In this issue of *Neuron*, Davidow et al. (2016) present a convincing argument that adolescent brain functionality can be adaptive in certain contexts, particularly probabilistic learning environments.

Adolescence is typically a time when children branch out, take risks, and seek reward. Reward-seeking behavior can refer to those activities that typically result in positive outcomes, such as trying a new sport or making new friends, or those that have the potential to put the adolescent in danger, such as illegal drug use (Spear, 2000). Many have suggested that reward-seeking behavior in adolescence can lead to troublesome outcomes and establish unhealthy habits (e.g., Steinberg, 2008), while others have suggested that these behaviors can be adaptive (e.g., Casey et al., 2010); empirical evidence for these two possibilities remains mixed. In the past few decades, new developments in technology such as fMRI have allowed scientists to examine what is happening in the brain when an adolescent makes a risky decision in a laboratory setting, such as during reinforcement learning.

Probabilistic learning refers to learning about associations that are true most, but not all, of the time. Learning about the associations occurs through trial and error, based on feedback on the learner's choices. This learning tends to be slow and often occurs outside of conscious awareness. For example, a classic laboratory task that involves probabilistic learning is the weather prediction task (Knowlton et al., 1994). In this task, participants are asked to look at playing cards with various symbols and learn to predict which combinations produce rain, and which produce sunshine. Some combinations of cards result in rain *most* of the time (say, 80%), while other combinations of cards result in sunshine most of the time. However, on 20% of trials where sunshine is normally produced, the outcome is rain. The participant must predict the most likely outcome, which sometimes means discounting occasional feedback that is incongruent with the majority of trials. In the same respect, in real life, just because the sky appeared sunny one day but produced rain does not mean you should carry an umbrella the next time you see clear skies.

During probabilistic learning and other feedback-based learning tasks, feedback results in a predictable signal change in a region within the brain's reward circuitry, known as the striatum (Delgado et al., 2005; Seymour et al., 2007). Feedback about correct and incorrect performance results in similar brain signals as those produced by extrinsic reward and punishments, such as monetary gain and loss, allowing behavior to be shaped to maximize future reward (Tricomi et al., 2006). These signals occur in response to a prediction error, which represents the difference between the actual outcome and the expected outcome. For example, when an actual outcome is more negative than the expected outcome, such as when a student expects an A grade but receives a C, this is termed a negative prediction error. A positive prediction error occurs when the opposite happens: the actual outcome is greater than expected, such as when a student expecting a C receives an A. Prediction error signaling in the ventral striatum and the hippocampus, an area vital to learning and memory, has been shown to aid in learning and memory (Sadeh et al., 2011), helping individuals learn from their previous successes and mistakes.

Most studies investigating reinforcement learning have focused on adult populations, but increasingly, research has suggested that differences in learning processes between adolescents and adults may help to explain why adolescents are more prone to engage in risk-seeking behavior than adults (Cohen et al., 2010; Galvan et al., 2006). Adolescents show significant brain differences compared to adults, both in terms of size of structures and in functionality (Steinberg, 2008; Luna et al., 2010). Many believe that delayed development of the prefrontal cortex, an area crucial to control and planning, results in differential use of major brain structures during adolescence as compared to adulthood (Steinberg, 2008; Casey et al., 2010; Luna et al., 2010). For example, during adolescence, areas of the brain that process reward develop earlier than the prefrontal cortex, which could explain adolescents' tendency toward reward-seeking behavior (Galvan et al., 2006). Although much of the prior research has focused on how these brain differences give rise to risky decision making that may lead to problematic outcomes (Galvan, 2013), the research presented by Davidow et al. (2016) suggests that these brain differences may actually result in adaptive strategies, such as during reward-based learning.

Davidow et al. (2016) examine whether reward sensitivity in adolescence improves learning when reinforcement is involved, as well as how differences in learning between adolescents and adults relate to interactions between the ventral striatum and the hippocampus. Using a probabilistic learning task, they tested whether adolescents (ages 13–17)

differed significantly from adults (ages 20–30) in terms of accuracy during learning, later episodic memory, and brain functionality. During fMRI data acquisition, participants were asked to pair different colored butterflies with various habitats. Similar to the weather prediction task, each butterfly was probabilistically associated with one habitat 80% of the time, and the other habitat 20% of the time. Responses were continuously reinforced by ''correct'' and ''incorrect'' feedback, so to succeed on the task, participants needed to update their choices based on previous feedback. Participants also saw images of objects while viewing feedback. After completing the task, participants completed a post-learning memory test about the butterfly task outside the scanner. In this episodic memory test, participants were shown the object images they originally viewed during feedback presentation, as well as new images, and then were asked to identify which photos they had already seen in the original task.

Davidow et al. (2016) found that the adolescents in their study showed better probabilistic learning performance than adults. Moreover, the authors also examined learning rates in their task, or the extent to which participants weighted new feedback to update future choices. Updating choices too quickly actually leads to suboptimal performance during this task, due to its probabilistic nature; too high of a learning rate would be like deciding that because it rained on what looked like a clear day yesterday, clear skies are now a sign that rain is imminent. Thus, a lower learning rate is more optimal because it discounts incongruous feedback. A slower learning rate suggests learning over a longer history of trials, rather than making predictions based only on the most recent trials. Davidow et al. (2016) found that the adolescents in their task exhibited lower learning rates. Additionally, although episodic memory in the adolescents was not significantly better than in adults, adolescents exhibited a significant relationship between prediction error and later memory: that is, prediction errors significantly related to stronger episodic memory for adolescents, while adults did not exhibit this relationship.

Davidow et al. (2016) also present evidence that suggests that differences between adolescents' and adults' learning may be a result of greater interaction between reinforcement learning and episodic memory processes in the adolescent brain. They found that both adults and adolescents showed correlations between behavioral prediction errors and activity in the striatum. However, adolescents exhibited significantly greater prediction error-related activity in the hippocampus compared to adults. The authors conducted a psychophysiological interaction (PPI) analysis to examine functional connectivity between the striatum and the hippocampus. When examining correct versus incorrect trials, Davidow et al. (2016) found greater connectivity between the putamen, a subregion within the striatum, and the hippocampus, but only in adolescents. They also found a significant correlation in adolescents between this connectivity and the extent to which positively reinforced events were remembered. Thus, the neural processes underlying episodic memory seem to better incorporate reward-related information in the adolescent than the adult brain.

Although many studies have focused on the negative aspects in which reward-seeking behavior and reward sensitivity plague adolescents, Davidow et al. (2016) present a convincing argument for when reward sensitivity in adolescence can have an adaptive side. Their results suggest that in a probabilistic learning setting, reward-sensitive adolescents can outperform adults. Their neural data also suggest that enhanced prediction error-related interactions between the hippocampus and striatum heighten memory for positively reinforced events. Their findings also suggest that the relationship between the striatum and hippocampus may have a stronger influence on learning during adolescence, which may enhance reward-seeking behavior, as well as reward-related memories. Future studies could examine whether and how underdeveloped prefrontal regions in adolescents alter interactions in this brain circuit, and how these interactions relate to behavioral flexibility and control. Another area of interest may be to examine whether participants who are more likely to make risky decisions in the real world are also better probabilistic learners in the laboratory.

Davidow et al. (2016)'s findings, however, should not be taken to mean that adolescents will necessarily be better at all types of reinforcement learning than adults. For example, adults seem to outperform adolescents on tasks that emphasize learning from negative feedback (Christakou et al., 2013). Further, whereas in a probabilistic learning task a lower learning rate is optimal, it remains unclear whether adolescents would outperform adults in a deterministic task, where the correct answer remains fixed, and therefore higher learning rates result in better performance (Ashby and Maddox, 2005). In this context, one wonders whether adolescents would also exhibit lower learning rates, or if they would be better than adults at adjusting their learning rates to meet the demands of the task. When one thinks of reward-seeking behavior in adolescence, one usually thinks of the ''sex, drugs, and rock and roll'' portrayed in Hollywood movies and not how risky behaviors are the result of reward-seeking mechanisms in the brain that may actually be adaptive. Although seeking both health-positive and health-negative reward during adolescence can relate to long-lasting habits and memories, one should consider how the adolescent brain processes reward when hoping to steer an adolescent in the ''most rewarding'' direction.

REFERENCES

Ashby, F.G., and Maddox, W.T. (2005). Annu. Rev. Psychol. *56*, 149–178.

Casey, B.J., Duhoux, S., and Malter Cohen, M. (2010). Neuron *67*, 749–760.

Christakou, A., Gershman, S.J., Niv, Y., Simmons, A., Brammer, M., and Rubia, K. (2013). J. Cogn. Neurosci. *25*, 1807–1823.

Cohen, J.R., Asarnow, R.F., Sabb, F.W., Bilder, R.M., Bookheimer, S.Y., Knowlton, B.J., and Poldrack, R.A. (2010). Nat. Neurosci. *13*, 669–671.

Davidow, J.Y., Foerde, K., Galván, A., and Shohamy, D. (2016). Neuron *92*, this issue, 93–99.

Delgado, M.R., Miller, M.M., Inati, S., and Phelps, E.A. (2005). Neuroimage *24*, 862–873.

Galvan, A. (2013). Curr. Dir. Psychol. Sci. *22*, 88–93.

Galvan, A., Hare, T.A., Parra, C.E., Penn, J., Voss, H., Glover, G., and Casey, B.J. (2006). J. Neurosci. *26*, 6885–6892.

Knowlton, B.J., Squire, L.R., and Gluck, M.A. (1994). Learn. Mem. *1*, 106–120.

Luna, B., Padmanabhan, A., and O'Hearn, K. (2010). Brain Cogn. *72*, 101–113.

Sadeh, T., Shohamy, D., Levy, D.R., Reggev, N., and Maril, A. (2011). J. Cogn. Neurosci. *23*, 1597–1608.

Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007). J. Neurosci. *27*, 4826–4831.

Spear, L.P. (2000). Neurosci. Biobehav. Rev. *24*, 417–463.

Steinberg, L. (2008). Dev. Rev. *28*, 78–106.

Tricomi, E., Delgado, M.R., McCandliss, B.D., McClelland, J.L., and Fiez, J.A. (2006). J. Cogn. Neurosci. *18*, 1029–1043.

Vision Restoration Becomes Druggable

Simon Darius Klapper[1] and Volker Busskamp[1,*]
[1]Center for Regenerative Therapies Dresden, TU Dresden, Fetscherstrasse 105, 01307 Dresden, Germany
*Correspondence: volker.busskamp@crt-dresden.de
http://dx.doi.org/10.1016/j.neuron.2016.09.030

In this issue of *Neuron*, Tochitsky et al. (2016) have identified the mechanism by which small-molecule photoswitches enter and specifically activate retinal OFF-ganglion cells in degenerated retinas. This drug development is a tremendous step toward the treatment of blindness, regardless of the underlying mutation.

The retina is the neural tissue at the back of the eye that captures the visual panorama. There are many hereditary diseases, such as retinitis pigmentosa (RP), that lead to blindness. RP is a genetically complex disease manifested by any one of more than 200 mutations. These mutations result in a common pathology, the progressive degeneration of the retina. Primarily rod photoreceptors, which are important for low-light vision, are affected by these mutations and degenerate. As the disease progresses, cone photoreceptors that mediate high acuity, daylight, and color vision become light insensitive and eventually also degenerate. At this stage of blindness, retinal remodeling occurs that progressively alters anatomical, biochemical, and functional features of the retina (Jones et al., 2016). There is no general therapeutic intervention available at present to restore visual function in RP.

Due to the high genetic complexity and the progressive nature of retinal degeneration, therapeutic interventions are challenging. The ideal option would be to substitute the mutated allele(s) by gene-replacement or gene-editing strategies prior to disease onset. Gene replacement involves the delivery of a wild-type copy to the affected cells to restore visual function and to prevent any further degeneration (Sahel and Roska, 2013). However, if a mutation is dominant, it is also necessary to silence the mutated gene product. Gene editing depends on the intrinsic DNA repair machinery, which limits gene correction by homolog-directed repair in postmitotic neurons; gene disruption via non-homologous end-joining, on the other hand, can be very efficient in retinal cells (Hung et al., 2016). A promising mutation-independent approach to restoring visual function is the application of optogenetic tools that confer light sensitivity to all targeted retinal cell types. In this way, and in the absence of intrinsic photoreceptors, targeted inner-retinal cell types become artificial photoreceptors (Klapper et al., 2016). A more invasive approach is the transplantation of electrical implants at late stages of RP, which results in the recovery of some visual function in RP patients (Chuang et al., 2014). Cell-replacement strategies aim to substitute lost photoreceptors or retinal pigment epithelium cells (Jayakody et al., 2015). Another promising option is to stop, or significantly delay, the onset and progression of the degeneration by applying neuroprotective factors that outpace the devastating effects of the primary mutations (Sahel and Roska, 2013). A combined functional and neuroprotective approach will likely be required to stop further retinal degeneration and restore visual perception.

In summary, there are many therapeutic interventions for RP currently under clinical evaluation or already being used to treat patients. However, these therapies depend on invasive surgeries, such as transplanting implants or injecting recombinant viral particles that carry a DNA-encoded cargo to targeted cell types in the eye.

Ideally, it would be better to have a small-molecule drug that restores visual function and can be either administered systemically or injected locally into the eye. Promising molecules have been described in pioneering studies; these photoisomerizable molecules bind to ion channels or receptors, thereby serving as photoswitches that render the endogenous proteins light sensitive. The photoisomerization can exert forces on the channels and receptors, as well as adding or removing ligands. These effects lead to channel activation and subsequently to neuronal activation when a certain depolarization threshold is reached (Kramer et al., 2009). These optochemical