

Journal Club

Editor's Note: These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml.

Dopaminergic Prediction Errors Persevere in the Nucleus Accumbens Core during Negative Reinforcement

Nick G. Hollon,^{1,2,3} Marta E. Soden,^{2,3} and Matthew J. Wanat^{2,3}

¹Graduate Program in Neurobiology and Behavior, ²Department of Psychiatry and Behavioral Sciences, and ³Department of Pharmacology, University of Washington, Seattle, Washington 98195

Review of Oleson et al.

One predominant theory of dopamine function postulates that phasic bursts and pauses in dopamine transmission signal errors in reward prediction, and can thereby adjust the probability of repeating a preceding action (Houk et al., 1995; Schultz et al., 1997). Specifically, dopaminergic prediction errors are thought to update the motivational value assigned to stimuli in the environment in a manner consistent with temporal difference reinforcement learning (TDRL), a prominent class of computational models used to explain learning about appetitive and aversive stimuli. During appetitive behavior, increases and decreases in dopamine transmission are observed when the value of an outcome is greater or less than expected, respectively. Furthermore, dopamine transmission evoked by reward-predictive stimuli can reflect changes in an animal's expectations of future reward. However, much less is known about dopamine signaling during negative reinforcement, i.e., strengthening a behavior

that prevents or removes an aversive outcome. A recent report by Oleson et al. (2012) has provided an important advance by examining phasic dopamine transmission during negative reinforcement.

Using fast-scan cyclic voltammetry to measure rapid changes in dopamine concentration in the core of the nucleus accumbens, Oleson et al. (2012) investigated mesolimbic dopamine responses to stimuli signaling the opportunity to avoid punishment by making an instrumental response. The avoidance learning task consisted of trials initiated by the simultaneous onset of a cue light and extension of a lever, which together served as a compound "warning signal." If a rat pressed the lever within 2 s of the warning signal onset, he avoided receiving any shock. Otherwise, the rat received repeated footshocks every 2 s until he eventually pressed the lever to make an escape response. After either avoidance or escape responses, a 20 s "safety period" was signaled by the offset of the cue light, retraction of the lever, and onset of a continuous auditory tone during which no shocks were delivered. The subsequent trial began with the offset of the safety period tone and the simultaneous onset of the warning signal. Dopamine transmission was recorded in a single voltammetry session after rats successfully avoided shock in at least 50% of trials for three consecutive behavioral sessions, allowing direct comparison of dopamine release during avoidance versus escape behaviors.

Oleson et al. (2012) measured an increase in dopamine release in response to the warning signal on trials in which rats successfully avoided footshock, but observed a decrease in dopamine transmission in response to the warning signal on trials in which rats failed to avoid footshock. They further suggested that dopamine concentration increases at the onset of the safety period following either avoidance or escape responses. In a separate Pavlovian fear conditioning experiment in which a conditioned stimulus (CS) signaled an unavoidable footshock, the authors found that this aversive CS reduced dopamine levels when rats froze during the CS, similar to the neurochemical response observed to the warning signal in instrumental escape trials when rats failed to avoid shock.

Oleson et al. (2012) suggest that the valence of the warning signal and the corresponding dopamine transmission change as the rats' behavior progresses from fear-induced responses, such as freezing, to the newly acquired avoidance behavior. The decreases in dopamine in response to the aversive Pavlovian CS and to the warning signal when rats fail to avoid shock could reflect the negative value assigned to these punishment-predictive stimuli. The authors interpret the increased dopamine in response to the warning signal on successful avoidance trials as analogous to the phasic increases also evoked by an appetitive CS (Schultz et al., 1997). Similarly, they claim that the

Received Dec. 17, 2012; revised Jan. 9, 2013; accepted Jan. 14, 2013.

N.G.H. was supported by a National Science Foundation graduate research fellowship, M.E.S. was supported by NIMH Grant R01 MH094536, and M.J.W. was supported by NIDA Grant R01 DA016782. We thank Drs. Paul Phillips, Jeremy Clark, Larry Zweifel, Jeansok Kim, and members of their labs for helpful comments during the preparation of this manuscript.

Correspondence should be addressed to Nick G. Hollon, Graduate Program in Neurobiology and Behavior, Department of Psychiatry and Behavioral Sciences, Box 356560, Health Sciences Building, 1959 NE Pacific Street, University of Washington, Seattle, WA 98195-6560. E-mail: nhollon@uw.edu.

DOI:10.1523/JNEUROSCI.5762-12.2013

Copyright © 2013 the authors 0270-6474/13/333253-03\$15.00/0

dopamine response to the onset of the safety period following either avoidance or escape behavior indicates that negative reinforcement increases dopamine release in a manner similar to unexpected positive reinforcement. From these collective findings, they propose that phasic dopamine transmission in the nucleus accumbens core indeed resembles a bidirectional prediction error signal that could also support negative reinforcement.

Oleson et al. (2012) interpret their results by citing Mowrer's (1947) "two-factor" theory and its modern TDRL-based adaptation using an "actor-critic" model (Maia, 2010). Both theories provide similar explanations for the acquisition of avoidance behavior: the warning signal initially acts as a Pavlovian CS that elicits an aversive motivational state of fear; any instrumental response that eliminates the warning signal and initiates a period of safety consequently would reduce this fear and would be reinforced. However, two-factor theory and TDRL differ in the role that reinforcement plays in maintaining the instrumental avoidance once animals reliably avoid shock on 100% of trials and the warning signal becomes less aversive as they receive no further punishment (Maia, 2010). According to two-factor theory, the animal's termination of the warning signal would provide little reduction of fear and thus would no longer be reinforcing, leading to the attenuation of avoidance behavior. In contrast, TDRL predicts no such extinction of avoidance behavior. With 100% avoidance, the fully expected absence of punishment would generate no prediction error at the time of safety period onset (avoidance expected, no punishment received). A prediction error equaling zero would neither strengthen nor weaken a previously acquired response, and avoidance behavior would be expected to persist. Because numerous empirical studies have reported the persistence of avoidance responses in the absence of further delivery of punishment, the actor-critic instantiation of TDRL appears better suited than two-factor theory to describe empirically observed avoidance behavior (Maia, 2010).

TDRL models can make specific predictions regarding dopamine transmission throughout the acquisition of instrumental avoidance behavior. Although Oleson et al. (2012) discuss their neurochemical results in terms of a progression throughout instrumental acquisition, they record dopamine release from only a single session in which each rat showed a mix of successful avoidance in

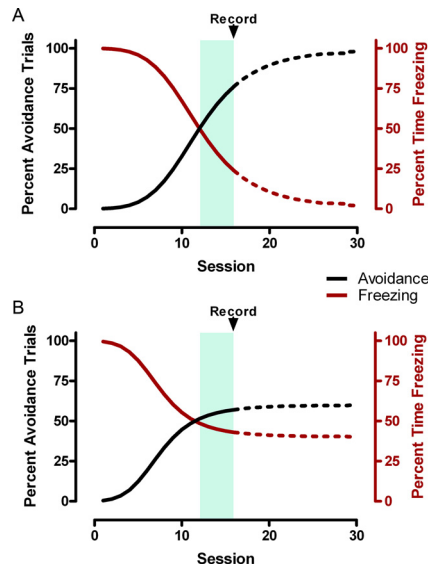


Figure 1. Hypothetical behavioral results extending those reported in Oleson et al. (2012), with black curves denoting percentage of trials with successful avoidance of footshock and red curves indicating percentage of time freezing during the warning signal. As described by Oleson et al. (2012), rats required ~15 sessions to reach a behavioral criterion of avoiding shock on at least 50% of trials for three consecutive sessions, indicated by the shaded region. After reaching this criterion, dopamine transmission in the nucleus accumbens core was recorded using fast-scan cyclic voltammetry in a single session, indicated by the arrow. Two possible behavioral outcomes are illustrated. **A**, The recorded session may have occurred at an intermediate stage of training as rats were still acquiring the avoidance response. With further training (dashed lines), they might eventually reach 100% avoidance and show little freezing behavior during the warning signal. **B**, Alternatively, the rats may have already reached approximately asymptotic behavior at just over 50% avoidance. Further training would not increase avoidance, and freezing might persist during the warning signal.

some trials and escape after failing to avoid shock in other trials. It is unclear whether this recorded session occurred after the rats had reached asymptotic avoidance or whether their performance on the task was still improving (Fig. 1). This behavioral distinction could prove critical in interpreting the neurochemical data and the hypothesized progression discussed by the authors. If we first assume that the rats' avoidance of shock on just over 50% of the trials indicates that they are still at an intermediate stage of acquiring the task (Fig. 1A), we might expect the pattern of dopamine transmission observed by Oleson et al. (2012) to change with further training. In theory, if the rats eventually avoid shock successfully on 100% of trials, a phasic increase in dopamine occurring upon the onset of the safety period would be a surprising result. After an animal learns reliably to avoid punishment, the prediction error at the onset of the safety

period approaches zero. Thus, if the rats were extensively trained to the extent that they avoided punishment on 100% of trials, one would expect to observe substantially attenuated dopamine release at the onset of the safety period upon successful shock avoidance, as is typically observed upon receipt of a fully expected reward (Schultz et al., 1997). If this were not the case, it would imply that dopamine signals during negative reinforcement follow different rules from those occurring during positive reinforcement.

Alternatively, it is possible that exhibiting just over 50% avoidance after ~15 training sessions reflects the rats' asymptotic level of performance in this particular task (Fig. 1B). Arbitrary responses such as lever pressing are notoriously more difficult for rats to acquire in instrumental avoidance tasks than are more natural responses such as fleeing to the other side of a shuttle box (Bolles, 1972). The use of manual shaping by successive approximations by Oleson et al. (2012) illustrates the challenge of training rats to make this particular instrumental response in this task. The failure to avoid shocks on any given trial might result from animals making species-specific defense reactions (Bolles, 1972); modern computational approaches to reinforcement learning have begun to consider how such innate behaviors might interact with acquired instrumental responses (Dayan et al., 2006). With asymptotic performance well below 100%, the dopaminergic responses to both the warning signal and safety period observed by Oleson et al. (2012) would be expected to persist throughout further training, as has been described for probabilistic rewards (Fiorillo et al., 2003). Conducting additional training as well as voltammetry recordings throughout avoidance learning would reveal whether the behavioral performance and corresponding dopamine transmission was indeed at an intermediate stage in learning (Fig. 1A) or already at an asymptotic level (Fig. 1B).

We also note that in the task used by Oleson et al. (2012), each trial involves several events occurring in close temporal proximity. Although the authors qualitatively describe the aforementioned increase in dopamine at the onset of the safety period, they present time courses of mean dopamine concentration aligned only to the onset of the warning signal. Because of the temporal jitter introduced by the variable response latencies, the neurochemical response to this safety signal cannot be isolated without realigning the voltammetry data to the time of the

instrumental response that initiates this safety period. A phasic increase during the onset of the safety period would represent the critical positive prediction error supporting negative reinforcement.

Additionally, it is difficult to derive an explanation from TDRL for why the warning signal would evoke an increase in dopamine transmission in successful avoidance trials. Unlike a reward-predictive stimulus, the warning signal in theory would not generate a positive prediction error. Specifically, as rats learn to reliably avoid shock, the expected future punishment would approach zero. Consequently, the value assigned to the warning signal and the temporal difference prediction error at that time would approach but never exceed zero. An alternative explanation for the observed increase in dopamine might be that it instead coincides with the initiation of the rats' approach toward the lever, which could also be assessed by realigning the voltammetry data to the instrumental response as suggested above (Roitman et al., 2004).

TDRL provides an informative framework with which to make predictions regarding dopamine transmission during positive and negative reinforcement. Of course, actual biology need not conform to any particular theoretical account, and recent work has illustrated the complexity of the dopaminergic system. For example,

subpopulations of putative dopamine neurons respond to aversive stimuli in a heterogeneous manner (Matsumoto and Hikosaka, 2009; Wang and Tsien, 2011). Indeed, Badrinarayan et al. (2012) recently replicated and extended the finding of Oleson et al. (2012) that an aversive Pavlovian CS rapidly decreased dopamine transmission in the nucleus accumbens core, but they also showed that this CS gradually increased dopamine levels in the nucleus accumbens shell. Future studies will undoubtedly continue to explore pathway-specific dopamine transmission during motivated behavior. Regardless, Oleson et al. (2012) have provided evidence that mesolimbic dopamine transmission in the nucleus accumbens core could be consistent with a reward-prediction error signal in negative reinforcement.

References

- Badrinarayan A, Wescott SA, Vander Weele CM, Saunders BT, Couturier BE, Maren S, Aragona BJ (2012) Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J Neurosci* 32:15779–15790. [CrossRef Medline](#)
- Bolles RC (1972) The avoidance learning problem. In: *The psychology of learning and motivation: advances in research and theory* (Bower GH, ed), pp 97–145. New York: Academic.
- Dayan P, Niv Y, Seymour B, Daw ND (2006) The misbehavior of value and the discipline of the will. *Neural Netw* 19:1153–1160. [CrossRef Medline](#)
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902. [CrossRef Medline](#)
- Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia* (Houk JC, Davis JL, Beiser DG, eds), pp 249–270. Cambridge, MA: MIT.
- Maia TV (2010) Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn Behav* 38:50–67. [CrossRef Medline](#)
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841. [CrossRef Medline](#)
- Mowrer OH (1947) On the dual nature of learning—a re-interpretation of “conditioning” and “problem-solving.” *Harv Educ Rev* 17:102–148.
- Oleson EB, Gentry RN, Chioma VC, Cheer JF (2012) Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J Neurosci* 32:14804–14808. [CrossRef Medline](#)
- Roitman MF, Stuber GD, Phillips PE, Wightman RM, Carelli RM (2004) Dopamine operates as a subsecond modulator of food seeking. *J Neurosci* 24:1265–1271. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Wang DV, Tsien JZ (2011) Convergent processing of both positive and negative motivational signals by the VTA dopamine neuronal populations. *PLoS One* 6:e17047. [CrossRef Medline](#)