
Jeff Beeler
 Departments of Neurobiology, Pharmacology, and Physiology, University of Chicago, Chicago, Illinois

We are all familiar with Pavlov’s dogs that saliva when the dinner bell rings or Skinner’s pigeons that learn that a peck on a key yields a delicious treat. But what about learning that is not focused on specific predictive cues or particular instrumental behaviors? What about the more amorphous type of learning that biases our behavior toward seeking a context previously associated with reward? How do we understand the drug addict drawn to people and places associated with drug use despite a desperate determination to remain sober? An important evolutionary challenge is learning to be in the right place at the right time; that is, learning where we are more likely to find reward and spending more time there even in the absence of specific predictive cues, like the dinner bell, or specific knowledge on how to make a reward occur. In this issue of the Journal of Neurophysiology (p. 2083–2093), German and Fields tackle this kind of learning by approaching a common behavioral paradigm from a novel perspective and in the process make exploring the mechanism of contextual reinforcement learning more tractable and intriguing for future research.

Consider the college freshman with an hour to kill between classes. He can either go to the coffee shop in the student union or the one in the library. On several occasions he has run into friends or struck up enjoyable conversations with strangers in the library canteen but rarely in the student union. Although he visits both venues throughout the quarter and often studies alone in both places, he is more likely to be found in the library coffee shop. On the Friday after Thanksgiving, he gets stuck on campus while all of his friends are away. He goes to the library that day, although he drinks his coffee alone. Conditioned place preference is an animal version of such contextual reinforcement learning. By focusing on the “where” of reinforcement learning, rather than when or how, this paradigm allows investigators to examine how reward related experience is associated with place—or context—and subsequently biases behavioral choices. In conditioned place preference, animals are placed in an apparatus divided into two (or 3) chambers. Preference for one chamber over the other(s) is assessed by measuring the amount of time spent in each chamber. During the conditioning phase, the opening between the chambers is blocked, and animals are exposed to each chamber separately while receiving a particular experience in each chamber. For example, the subject may receive morphine in the left chamber and saline in the right. During the preference test, the animal is allowed to roam freely between the chambers, and no experience (e.g., morphine or saline) is provided. Thus the animal’s preference during the test phase is based solely on previous associations with each chamber.

Conditioned place preference, as widely used, focuses on the acquisition of preference (frequently assessing the motivational value of drugs) while the discrete behaviors that comprise the expression of preference are generally unexamined (Tzschentke 1998). German and Fields reframe the conditioned place preference task as an exploratory decision-making process for the rodent. At any given moment the animal has to decide whether to stay put or to move, and if it moves, it must decide where to go. “Preference” arises from a series of behavioral choices that reflect an exploratory strategy. Examining the duration of visits in each chamber and transitions between chambers, the authors found that the animals do not favor the morphine-paired room with longer visits (which would be observed as a shifted Gaussian distribution of visit duration), but rather they exhibit an increased exit probability (a shifted exponential distribution of exits) from the saline-paired room. This is curious in the context of addiction and relapse. The rats are not merely associating the morphine-paired chamber with good times and lingering in the chamber but are actively decreasing their exploration of the saline-paired room and thereby, given the restricted choices, increasing their rate of return to the morphine-paired chamber. When the previous experience does not occur in the morphine-paired room, however, they continue to explore. As German and Fields put it “they have reason to go, but no reason to stay.” This reflects the experience of an addict during early sobriety: nondrug-related cues and contexts are not rewarding but neither are drug-related cues and contexts in the absence of drug. What is counterintuitive is that the actual change in goal-directed, exploratory behavior occurs in the previously nonrewarded context. This finding is provocative and suggests an alternative to conventional thinking about relapse: an important component of drug-seeking behavior during abstinence may be an increased propensity to “give up” on nonrewarded contexts; that is, a decrease in tolerance for and exploration of previously nonrewarded environments may be a salient underlying behavioral process promoting relapse.

In conditioned place preference, there are no distinct, temporally discrete cues predicting reward, nor are there any instrumental behaviors with which the animal brings about reward. In fact, during the testing phase, there is no reward. The measured behavior, preference, is an aggregate statistic rather than a discrete response. Therefore there is no cue, learned response, or reward consumption within which to frame a standard electrophysiological analysis. The model described by German and Fields overcomes this difficulty and offers a discrete behavior to be examined.

Address for reprint requests and other correspondence: J. Beeler (E-mail: jabeeler@uchicago.edu).

First published January 10, 2007; doi:10.1152/jn.01318.2006.
Conditioned place preference is often construed as a form of Pavlovian conditioning where the CS is a complex compound stimulus (i.e., context) and the CR is simply approach behavior. Although theoretically satisfying, this view obscures potentially interesting questions about place learning (see Bardo and Bevins 2000). How is the compound CS represented in the brain? Are the various cues comprising a “place” coalesced into a configuration of stimuli and encoded by specific cells representing context? In the absence of a discrete CR, such as salivating, what response is associated with the CS (see Konorski 1966)?

In a companion paper in this issue of the *Journal of Neurophysiology* (p. 2094–2106), German and Fields enter into this largely uncharted territory. The nucleus accumbens (NAc) plays a critical role in conditioned place preference (Bardo 1998). By examining the correlation of NAc activity with transitions between rooms before and after conditioning, German et al. provide a first glimpse of the mechanics of the accumbens’ contribution to place conditioning. They find a population of neurons that exhibit room specific activity after conditioning, including neurons excited by the saline-paired room and inhibited by the morphine-paired room. The activities associated with each room are tonic, sustained throughout the duration of a room visit, and change immediately on transition to a different location. Thus this activity constitutes not a direct motor signal preceding transitions but a bias operating during a visit, which the authors argue determines the probability of exiting from the chamber—a “context-dependent preference signal.” Far from resolving the mechanism of conditioned place preference, this first direct look at the underlying neural activity promises to open Pandora’s box and stimulate both debate and future studies.

The apparent simplicity of the subjective experience of reward belies the complexity of reward and reward learning as neural processes. Recent computational models have enriched our theoretical understanding of reinforcement learning (Montague et al. 2004). German and Fields present a model of the rat’s behavior in conditioned place preference that accurately describes and simulates their actual behavior; it does not, however, explain the phenomenon it describes. Their model will stand as an invitation to bring the same theoretical rigor and formal computational investigation to conditioned place preference as has been afforded other learning tasks, potentially yielding a richer, more comprehensive understanding of reinforcement learning generally. Although the initial insight underlying these two papers appears humble enough—asking what specific behaviors comprise preference—this simple beginning promises to open fertile new areas of investigation.

REFERENCES


