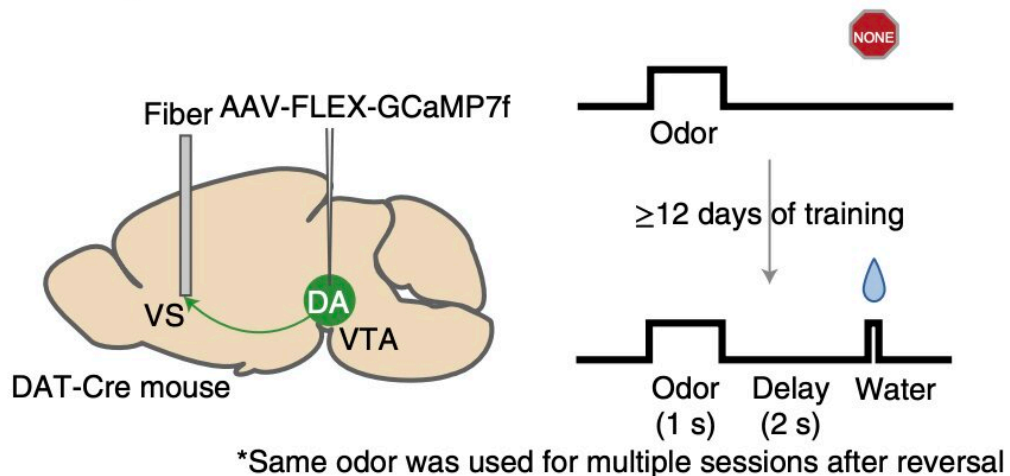


Study draws new link between dopamine-based reward learning and machine learning

July 27 2022, by Ingrid Fadelli

Reversal learning



Credit: Amo et al.

Past neuroscience and psychology research has repeatedly demonstrated the crucial role of rewards in how humans and other animals acquire behaviors that promote their survival. Dopaminergic neurons, neurons that produce dopamine in the mammalian central nervous system, are known to be largely responsible for reward-based learning in mammals.

Studies have found that when a mammal is receiving an unexpected reward, these neurons promptly respond, through a so-called phasic excitation. This is a short but strong period of excitation that occurs in rapidly adapting [brain receptors](#) (i.e., phasic receptors).

When animals start creating associations between a reward and a specific stimulus or cue, dopamine neurons tune down their responses to the reward in question. This might be an evolutionary mechanism aimed at supporting associative learning.

In recent years, computer scientists have been trying to artificially replicate the neural underpinnings of reward-learning in mammals, to create efficient machine learning models that can tackle [complex tasks](#). A renowned machine learning method that replicates the function of [dopaminergic neurons](#) is the so-called temporal difference (TD) learning algorithm.

Researchers at Harvard University, Nagoya University and Keio University School of Medicine have recently carried out a study exploring an aspect of the TD learning computational method that could be related to how humans learn based on rewards. Their paper, published in *Nature Neuroscience*, could shed some new light on how the brain builds associations between cues and rewards that are separated in time (i.e., that are not consecutive and thus not experienced one after another).

TD learning algorithms are a class of reinforcement learning approaches that do not require a model, but rather can learn to make predictions based on changes in the environment that occur over successive time steps. In contrast with other machine learning methods, TD methods can adjust their estimations several times before they reveal their final prediction.

Over the past few years, several studies have highlighted the similarities between TD learning algorithms and reward-learning dopamine neurons in the brain. Nonetheless, one particular aspect of the algorithm's functioning has rarely been considered in neuroscience research.

"Previous studies failed to observe a key prediction of this algorithm: that when an agent associates a cue and reward that are separated in time, the timing of dopamine signals should gradually move backward in time from the time of the reward to the time of the cue over multiple trials," Ryunosuke Amo, Sara Matias, Akihiro Yamanaka, Kenji F. Tanaka, Naoshige Uchida and Mitsuko Watabe-Uchida wrote in their paper. "We demonstrate that such a gradual shift occurs both at the level of dopaminergic cellular activity and dopamine release in the ventral striatum in mice."

In their paper, Amo and his colleagues considered the results of experiments they carried out on untrained mice who were learning to associate odor cues with water rewards. When the animals started associating specific odors with receiving water, they exhibited licking behavior suggesting that they were expecting to receive water after they had only smelled the associated odor.

In their experiments, the researchers presented the mice with the pre-reward odor and the reward within different time frames. In other words, they changed the amount of time between the moment when the mice were exposed to the odor and the moment when they received the water reward.

They found that when the reward was delayed (i.e., it was given to the mice later than they had previously experienced), [dopamine neurons](#) were not so active in the beginning but became more active as time went on. This showed that, as observed in TD learning approaches, the timing of dopamine responses in the brain can shift while mice are learning

associations between odors and rewards for the first time.

The team also conducted further experiments, to test whether this shift also occurred in animals that had already been trained to make these odor-reward associations and during reverse tasks (i.e., tasks where the cue and reward were reversed). They observed a temporal shift in the animal's dopamine signals during the delay period, which was similar to that exhibited when animals were learning associations for the first time, but at a faster speed.

Overall, the findings gathered by Amo and his colleagues highlight the occurrence of a backward shift in the timing of [dopamine](#) activity in the mice brain throughout different associative learning experiments. This observed temporal shift greatly resembles the mechanisms underpinning TD learning methods.

In the future, the findings gathered by this team of researchers could pave the way for new studies investigating this potential similarity between reward-learning in the mammalian brain and TD reinforcement learning approaches. This could help to improve the current understanding of reward-learning in the brain, while also potentially inspiring the further development of TD learning algorithms.

"Our results establish a long-sought link between dopaminergic activity and the TD learning algorithm, providing fundamental insights into how the brain associates cues and rewards that are separated in time," the researchers wrote in their paper.

More information: Ryunosuke Amo et al, A gradual temporal shift of dopamine responses mirrors the progression of temporal difference error in machine learning, *Nature Neuroscience* (2022). [DOI: 10.1038/s41593-022-01109-2](https://doi.org/10.1038/s41593-022-01109-2)

© 2022 Science X Network

Citation: Study draws new link between dopamine-based reward learning and machine learning (2022, July 27) retrieved 26 April 2024 from <https://medicalxpress.com/news/2022-07-link-dopamine-based-reward-machine.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.