Article

Current Biology

# Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision

## Highlights

- Reinforcement learning model with belief state to cope with perceptual uncertainty
- Model provides unified account of dopamine in perceptual and reward-guided choices
- Dopamine can act as a teaching signal during perceptual decision making as well
- Dopamine signals decision confidence prior to behavioral manifestation of choice

## Authors

Armin Lak, Kensaku Nomoto, Mehdi Keramati, Masamichi Sakagami, Adam Kepecs

## Correspondence

kepecs@cshl.edu

## In Brief

Lak et al. show that dopamine neuron responses during a visual decision task comply with predictions of a reinforcement learning model with a belief state signaling confidence. The results reveal that dopamine neurons encode teaching signals appropriate for learning perceptual decisions and respond early enough to impact impending decisions.

Lak et al., 2017, Current Biology *27*, 1–12
March 20, 2017 
http://dx.doi.org/10.1016/j.cub.2017.02.026

CellPress

Please cite this article in press as: Lak et al., Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision, Current Biology (2017), http://dx.doi.org/10.1016/j.cub.2017.02.026

Current Biology
Article



# Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision

Armin Lak,[1,2] Kensaku Nomoto,[3,4] Mehdi Keramati,[5] Masamichi Sakagami,[3] and Adam Kepecs[1,6,*]
[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[2]UCL Institute of Ophthalmology, University College London, London EC1V 9EL, UK
[3]Brain Science Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan
[4]Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Av. de Brasilia, 1400-038 Lisbon, Portugal
[5]Gatsby Computational Neuroscience Unit, University College London, London W1T 4JG, UK
[6]Lead Contact
*Correspondence: kepecs@cshl.edu
http://dx.doi.org/10.1016/j.cub.2017.02.026

## SUMMARY

**Central to the organization of behavior is the ability to predict the values of outcomes to guide choices. The accuracy of such predictions is honed by a teaching signal that indicates how incorrect a prediction was ("reward prediction error," RPE). In several reinforcement learning contexts, such as Pavlovian conditioning and decisions guided by reward history, this RPE signal is provided by midbrain dopamine neurons. In many situations, however, the stimuli predictive of outcomes are perceptually ambiguous. Perceptual uncertainty is known to influence choices, but it has been unclear whether or how dopamine neurons factor it into their teaching signal. To cope with uncertainty, we extended a reinforcement learning model with a belief state about the perceptually ambiguous stimulus; this model generates an estimate of the probability of choice correctness, termed decision confidence. We show that dopamine responses in monkeys performing a perceptually ambiguous decision task comply with the model's predictions. Consequently, dopamine responses did not simply reflect a stimulus' average expected reward value but were predictive of the trial-to-trial fluctuations in perceptual accuracy. These confidence-dependent dopamine responses emerged *prior to* monkeys' choice initiation, raising the possibility that dopamine impacts impending decisions, in addition to encoding a post-decision teaching signal. Finally, by manipulating reward size, we found that dopamine neurons reflect both the upcoming reward size and the confidence in achieving it. Together, our results show that dopamine responses convey teaching signals that are also appropriate for perceptual decisions.**

## INTRODUCTION

In the struggle of life, animals survive by following a simple dictum: win big and win often [1]. Finding bigger wins (e.g., more food reward) and more likely wins is particularly challenging when these are not available in their nearby environment. In these situations, a process of trial and error is required to selectively reinforce the most successful actions. Inspired by the study of animal behavior, a machine learning approach called reinforcement learning provides a rigorous framework to understand how to select winning behaviors. The key to reinforcement learning is adjusting the expected reward values associated with each behavior based on the *outcomes* of one's actions. These adjustments to reward values are based on the discrepancy between the received and predicted value, referred to as the reward prediction error [2]. There is a great deal understood about the neural mechanisms underlying reinforcement learning, and it is well established that midbrain dopamine neurons broadcast reward prediction error signals [3–6]. Here we address whether dopamine neurons provide appropriate prediction error signals when there is ambiguity in the cues that predict reward.

Computing reward prediction error, by definition, requires predicting the value of impending outcomes. Such value prediction relies on different sources of information and correspondingly distinct processes as dictated by the behavioral context. In one context, distinct, unambiguous cues that predict different reward outcomes are used to guide decisions. Because there is no uncertainty in identifying the cues, the accuracy of outcome predictions is limited instead by potentially complex, probabilistic reward payoff contingencies. Thus, the expected value of each decision can be estimated based on the experienced outcomes associated with the cues. These estimates can be produced by classic reinforcement learning algorithms [2]. In the context of ambiguous stimuli requiring perceptual decisions, animals face an additional challenge, because reward history alone can only provide an inaccurate estimate of upcoming outcome value. Rather, estimating the value of the choice requires an evaluation of the immediate percept and the decision process to compute the probability that the choice will be correct [7–9]. Thus, reward-history-guided and perceptual decisions, despite having fundamental similarities, differ in the computations required for reward prediction and hence prediction error estimation.

The phasic activity of dopamine neurons has been the subject of many studies, a few employing choice behaviors and many using simple Pavlovian conditioning tasks [10, 11]. The results of these studies can be chiefly summarized as showing that