The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty

Highlights

- Dopamine reward prediction errors (RPEs) reflect hidden state inference
- Medial prefrontal cortex (mPFC) shapes RPEs in a task involving hidden states
- mPFC is not needed to compute RPEs in a similar task when states are fully observable
- Modeling suggests that mPFC computes a probability distribution over hidden states

Authors

Clara Kwon Starkweather, Samuel J. Gershman, Naoshige Uchida

Correspondence

gershman@fas.harvard.edu (S.J.G.), uchida@mc.b.harvard.edu (N.U.)

In Brief

Dopamine neurons signal reward prediction errors, driving reinforcement learning. In ambiguous settings, dopamine signals incorporate hidden state inference. We demonstrate that the medial prefrontal cortex is required for hidden state inference to influence dopamine signals, illuminating the neural circuit governing reinforcement learning under state uncertainty.
The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty

Clara Kwon Starkweather,1 Samuel J. Gershman,2,* and Naoshige Uchida1,3,*
1Center for Brain Science, Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA
2Center for Brain Science, Department of Psychology, Harvard University, 52 Oxford Street, Cambridge, MA 02138, USA
3Lead Contact
*Correspondence: gershman@fas.harvard.edu (S.J.G.), uchida@mcb.harvard.edu (N.U.)
https://doi.org/10.1016/j.neuron.2018.03.036

SUMMARY

Animals make predictions based on currently available information. In natural settings, sensory cues may not reveal complete information, requiring the animal to infer the “hidden state” of the environment. The brain structures important in hidden state inference remain unknown. A previous study showed that midbrain dopamine neurons exhibit distinct response patterns depending on whether reward is delivered in 100% (task 1) or 90% of trials (task 2) in a classical conditioning task. Here we found that inactivation of the medial prefrontal cortex (mPFC) affected dopaminergic signaling in task 2, in which the hidden state must be inferred (“will reward come or not?”), but not in task 1, where the state was known with certainty. Computational modeling suggests that the effects of inactivation are best explained by a circuit in which the mPFC conveys inference over hidden states to the dopamine system.

INTRODUCTION

The ability to predict future outcomes is at the core of adaptive behaviors. In reinforcement learning theories, future outcomes are predicted based on the “state” of the world defined by a set of information, including the location of the animal, what objects are present, and elapsed time from certain events. A challenge in making predictions in natural environments is that the cues required to define the current state are often ambiguous, and it is difficult to know in which state the animal is in the first place. That is, the current state is “hidden” and needs to be inferred from partial information (Courville et al., 2006; Gershman et al., 2010, 2015). It has been proposed that, in the presence of such uncertainty, the brain computes future expectations based on a probability distribution defined over possible hidden states (a “belief state”; Daw et al., 2006; Rao, 2010). Although empirical evidence has begun to support this idea as it applies to the midbrain dopamine system (Rao, 2010; Lak et al., 2017; Starkweather et al., 2017), the neural mechanisms underlying hidden state inference remain largely unknown.

The activity of dopamine neurons is sensitive to reward expectation. Dopamine neurons report a reward prediction error (RPE) signal thought to reflect the discrepancy between actual and predicted value (Schultz et al., 1997; Bayer and Glimcher, 2005; Cohen et al., 2012; Eshel et al., 2015). Dopamine neurons’ responses to reward-predictive cues scale with the expected future reward (Fiorillo et al., 2003; Cohen et al., 2012; Tian and Uchida, 2015). More importantly, dopamine reward responses are suppressed according to the magnitude of reward expectation (Fiorillo, Tobler, and Schultz, 2003; Cohen et al., 2012; Tian and Uchida, 2015). The magnitude of dopamine reward responses is modulated by the moment-by-moment strength of reward expectation when the timing of the reward is varied (Fiorillo et al., 2008; Nomoto et al., 2010; Pasquereau and Turner, 2015). For instance, dopamine reward responses decrease as time elapses, as if reward expectation increases as a function of elapsed time. This result is consistent with the idea that reward expectation grows with the hazard rate (i.e., the likelihood of an event happening at a moment, given that the event has not happened yet). Notably, these prior studies with variable reward delivery times have utilized experimental paradigms in which reward is always delivered. In contrast, a previous study showed that the hazard account does not hold under conditions in which the reward is delivered in a probabilistic manner, and, instead, a model incorporating hidden state inference explains the data better (Starkweather et al., 2017).

This previous study (Starkweather et al., 2017) recorded dopamine RPEs during a classical conditioning task in which reward timing was varied across trials (Figure 1). The activity of dopamine neurons exhibited distinct patterns of responses depending on whether the reward was delivered in 100% of trials (task 1) or 90% of trials (task 2). In task 1, dopamine reward responses were modulated negatively over time, as if expectation increased over time, consistent with the hazard rate account (Figure 1D). In a stark contrast, in task 2, dopamine reward responses increased as time elapsed (Figure 1H). Computational modeling in which reward expectation is computed over belief states explained these divergent patterns. This model assumes transitions between two states: the inter-stimulus interval (ISI) state, during which reward is expected, and the inter-trial interval (ITI) state, during which no reward is expected (Figures 1B and 1F). The animal infers which state it is in based on the presentation of odor cues, reward, and the elapsed time from these events. Importantly, probabilistic reward delivery renders the task states hidden; the animal cannot know for certain whether
دریافت فوری

<table>
<thead>
<tr>
<th>متن کامل مقاله</th>
</tr>
</thead>
</table>

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات