Introduction

To successfully learn from reward feedback, the brain must adjust how it responds to and integrates reward outcomes, since reward contingencies can unpredictably change over time [1,2]. At the heart of this learning problem is a tradeoff between adaptability and precision. On the one hand, the brain must rapidly update reward values in response to changes in the environment; on the other hand, in the absence of any such changes, it must obtain accurate estimates of those values. This tradeoff, which we refer to as the adaptability-precision tradeoff [3,4], can be easily demonstrated in the framework of reinforcement learning [5]. According to this framework, larger learning rates result in higher adaptability but lower precision, and smaller learning rates give rise to lower adaptability but higher precision. In recent years, the failure of conventional reinforcement learning (RL) models to capture the level of adaptability and precision demonstrated by humans and animals has led to alternative explanations for how we deal with uncertainty and volatility in the environment [1,6,7,8]. However, most of these solutions for adjusting learning require complicated calculations, and their underlying neural substrates are unknown.

Given the central role of synapses in learning, we asked whether there are local synaptic mechanisms that can adjust the level of plasticity according to reward statistics and, therefore, allow the learning process to be adaptable. A candidate mechanism for such adjustment is metaplasticity, defined as changes in the synaptic state that shape the direction, magnitude, and duration of future synaptic changes without any observable change in the efficacy of synaptic transmission [9,10,11,12]. Extending our recent heuristic model of reward-dependent metaplasticity, which enables adjustment of learning to reward uncertainty [3], we examined a general class of metaplastic models to identify features that are beneficial for mitigating the adaptability-precision tradeoff (APT) during the estimation of the probability of binary reward.

Using the mean-field and Monte Carlo simulations, we identified optimal metaplastic models that can substantially overcome the APT. These models, which we refer to as 'superior' models, achieve both adaptability and precision by forming two separate sets of meta-states: reservoirs and buffers. Synapses in reservoir meta-states do not change their efficacy upon reward feedback, whereas those in buffer meta-states can change their efficacy. In superior models, rapid changes in efficacy are limited to synapses occupying buffers, creating a bottleneck that reduces noise without significantly decreasing adaptability. In contrast, more-populated reservoirs can generate a strong signal without manifesting any observable plasticity. Comparison of the behavior of our model and a few competing models during a dynamic probability estimation task revealed that superior metaplastic models perform close to optimally for a wider range of model parameters. However, superior models were suboptimal when precision was defined in the absolute term (absolute value of the difference between estimated and actual probabilities) rather than relative (the ability to distinguish neighboring values of probability), indicating that the brain could use different objectives to deal with reward uncertainty. Finally, we showed that metaplasticity provides a robust mechanism for mitigating the APT, and that metaplastic transitions are crucial, since replacing these transitions with plastic ones reduces the ability of the model to mitigate the APT. Altogether, our results illustrate how metaplasticity can mitigate one of the most fundamental tradeoffs in learning and, moreover, reveal the critical features of metaplasticity that contribute to adaptive learning.

Results