**Abstract**

Confidence estimation is a crucial component of decision making. Plenty of studies on decision confidence and uncertainty have focused on model-free value-based perceptual decision making. However, there has been less of a focus on confidence judgements in model-based learning and decision making in which animal forms a mental model of the environment to guide planning and adaptive behavior. In this work, we explore how animals estimate the uncertainty and develop decision confidence using deep reinforcement learning agents. We extended the post-decision wagering paradigm that is used to study decision confidence to study model confidence in a modified version of the two-step task. Specifically, the agent is asked to predict and report an upcoming state transition and receive a reward or penalty based on the correctness of model judgement. Crucially, the agent could also choose a third safe-bet option that gives them a small fraction of positive reward regardless of the correctness of their model judgements. Furthermore, at the start of each trial, the agent could choose to either perform the current trial or skip the current trial entirely and obtain a safe-bet reward. We expect our trained agent to choose the safe-bet option more when they are less confident about their learnt model. In addition, we hypothesize that agents’ model confidence in later stages of a trial predicts model confidence in earlier stages of a trial, giving a hierarchical structure of model confidence and decision making.