Researchers develop metrics to quantify information in animal responses to reward feedback

November 16 2021

Credit: Unsplash/CC0 Public Domain

Daily life is full of choices. How humans and other animals decide to allocate tiime and effort across competing priorities has fascinated

researchers for decades. Psychologists have found that most animals allocate their time among options in proportion to rewards received from the options, adjusting their behavior accordingly in response to reward feedback, a behavioral law known as "matching." A Dartmouth-led research team has developed metrics that measure the information content of response to reward feedback and used these metrics to predict global matching behavior. The results are published in *Nature Communications*.

"Matching is fundamental to how we choose between options available to us," says co-first author Ethan Trepka '22, a student in the Computational and Cognitive Neuroscience Lab at Dartmouth. "It governs things like which checkout line we choose at the grocery store or how much time we spend on different projects for school or work. How much time one chooses to spend on a given option depends on how frequently a reward is received from that option relative to other options."

In collaboration with other researchers, a Dartmouth team re-analyzed behavioral data from mice and monkeys collected at Johns Hopkins University and the National Institutes of Health, respectively. In the experiments, mice and monkeys chose between two options or actions and received rewards—water for mice and drops of apple juice for monkeys—based on their choices. The option with the better reward probability could change, so the animals had to keep track of the previous rewards as they made their choices. The results showed that both mice and monkeys exhibited undermatching, a general tendency to select the better option less than what is prescribed by the matching law.

To predict this deviation from the matching law, the researchers developed a new set of metrics that measure "surprise" or inconsistency in the animal's tendency to stay on or switch from the current option depending on the reward outcome. The metrics are based on the concept

of "entropy" in "information theory," a mathematical framework that can be used to quantify the amount of uncertainty or surprise in a system. Prior research on choice behavior has typically relied on computational models that require assumptions be made about what is learned from each reward feedback but the new metrics do not require such assumptions. The metrics also provide a new way to quantify adaptive behavior and can be used to improve previous computational models of learning and decision making.

"When we are faced with different options, we use the outcomes of our previous choices to make future decisions, and this should make us choose the better (more rewarding) option most of the time," says senior author Alireza Soltani, an associate professor of psychological and brain sciences and principal investigator of the Computational and Cognitive Neuroscience Lab at Dartmouth. "However, we do not choose the better option as often as we should and end up undermatching. Although choosing inferior options allows for new opportunities to be discovered in changing environments, when reward feedback on inferior options is disregarded and the response is inconsistent, undermatching increases significantly. Since undermatching often reduces the total reward that can be obtained, it is therefore, considered undesirable."

More information: Ethan Trepka et al, Entropy-based metrics for predicting choice behavior based on local response to reward, *Nature Communications* (2021). DOI: 10.1038/s41467-021-26784-w