Neuroscientists find brain activity patterns that encode our beliefs

July 15 2019

Credit: CC0 Public Domain

For decades, research has shown that our perception of the world is influenced by our expectations. These expectations, also called "prior beliefs," help us make sense of what we are perceiving in the present,

based on similar past experiences. Consider, for instance, how a shadow on a patient's X-ray image, easily missed by a less experienced intern, jumps out at a seasoned physician. The physician's prior experience helps her arrive at the most probable interpretation of a weak signal.

The process of combining prior knowledge with uncertain evidence is known as Bayesian integration and is believed to impact widely our perceptions, thoughts, and actions. Now, MIT neuroscientists have discovered distinctive brain signals that encode these prior beliefs. They have also found how the brain uses these signals to make judicious decisions in the face of uncertainty.

"How these beliefs come to influence brain activity and bias our perceptions was the question we wanted to answer," says Mehrdad Jazayeri, the Robert A. Swanson Career Development Professor of Life Sciences, a member of MIT's McGovern Institute for Brain Research, and the senior author of the study.

The researchers trained animals to perform a timing task in which they had to reproduce different time intervals. Performing this task is challenging because our sense of time is imperfect and can go too fast or too slow. However, when intervals are consistently within a fixed range, the best strategy is to bias responses toward the middle of the range. This is exactly what animals did. Moreover, recording from neurons in the frontal cortex revealed a simple mechanism for Bayesian integration: Prior experience warped the representation of time in the brain so that patterns of neural activity associated with different intervals were biased toward those that were within the expected range.

MIT postdoc Hansem Sohn, former postdoc Devika Narain, and graduate student Nicolas Meirhaeghe are the lead authors of the study, which appears in the July 15 issue of *Neuron*.

Ready, set, go

Statisticians have known for centuries that Bayesian integration is the optimal strategy for handling uncertain information. When we are uncertain about something, we automatically rely on our prior experiences to optimize behavior.

"If you can't quite tell what something is, but from your prior experience you have some expectation of what it ought to be, then you will use that information to guide your judgment," Jazayeri says. "We do this all the time."

In this new study, Jazayeri and his team wanted to understand how the brain encodes prior beliefs, and put those beliefs to use in the control of behavior. To that end, the researchers trained animals to reproduce a time interval, using a task called "ready-set-go." In this task, animals measure the time between two flashes of light ("ready" and "set") and then generate a "go" signal by making a delayed response after the same amount of time has elapsed.

They trained the animals to perform this task in two contexts. In the "Short" scenario, intervals varied between 480 and 800 milliseconds, and in the "Long" context, intervals were between 800 and 1,200 milliseconds. At the beginning of the task, the animals were given the information about the context (via a visual cue), and therefore knew to expect intervals from either the shorter or longer range.

Jazayeri had previously shown that humans performing this task tend to bias their responses toward the middle of the range. Here, they found that animals do the same. For example, if animals believed the interval would be short, and were given an interval of 800 milliseconds, the interval they produced was a little shorter than 800 milliseconds. Conversely, if they believed it would be longer, and were given the same

800-millisecond interval, they produced an interval a bit longer than 800 milliseconds.

"Trials that were identical in almost every possible way, except the animal's [belief](link) led to different behaviors," Jazayeri says. "That was compelling experimental evidence that the animal is relying on its own belief."

Once they had established that the animals relied on their prior beliefs, the researchers set out to find how the brain encodes prior beliefs to guide behavior. They recorded activity from about 1,400 neurons in a region of the frontal cortex, which they have previously shown is involved in timing.

During the "ready-set" epoch, the activity profile of each neuron evolved in its own way, and about 60 percent of the neurons had different activity patterns depending on the context (Short versus Long). To make sense of these signals, the researchers analyzed the evolution of neural activity across the entire population over time, and found that prior beliefs bias behavioral responses by warping the neural representation of time toward the middle of the expected range.

Embedded knowledge

Researchers believe that prior experiences change the strength of connections between neurons. The strength of these connections, also known as synapses, determines how neurons act upon one another and constrains the patterns of activity that a network of interconnected neurons can generate. The finding that prior experiences warp the patterns of neural activity provides a window onto how experience alters synaptic connections. "The brain seems to embed prior experiences into synaptic connections so that patterns of brain activity are appropriately biased," Jazayeri says.

As an independent test of these ideas, the researchers developed a computer model consisting of a network of neurons that could perform the same ready-set-go task. Using techniques borrowed from machine learning, they were able to modify the synaptic connections and create a model that behaved like the animals.

These models are extremely valuable as they provide a substrate for the detailed analysis of the underlying mechanisms, a procedure that is known as "reverse-engineering." Remarkably, reverse-engineering the model revealed that it solved the task the same way the monkeys' brain did. The model also had a warped representation of time according to prior experience.

The researchers used the computer model to further dissect the underlying mechanisms using perturbation experiments that are currently impossible to do in the brain. Using this approach, they were able to show that unwarping the neural representations removes the bias in the behavior. This important finding validated the critical role of warping in Bayesian integration of prior knowledge.

The researchers now plan to study how the brain builds up and slowly fine-tunes the synaptic connections that encode prior beliefs as an animal is learning to perform the timing task.

Provided by Massachusetts Institute of Technology