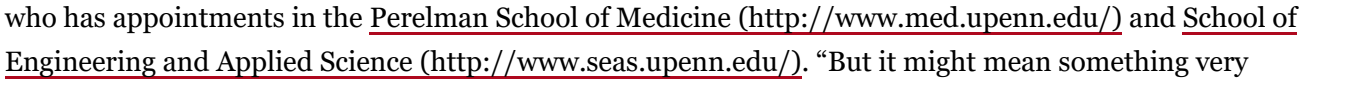
SCIENCE & TECHNOLOGY

Defining neural 'representation'

Neuroscientists frequently say that neural activity 'represents' certain phenomena. PIK Professor Konrad Kording and postdoc Ben Baker led a study that took a philosophical approach to tease out what the term means.



Neuroscientists use the word "represent" to encompass multifaceted relationships between brain activity, behavior, and the environment.

One of neuroscience's greatest challenges is to bridge the gaps between the external environment, the brain's internal electrical activity, and the abstract workings of behavior and cognition. Many neuroscientists rely on the word "representation" to connect these phenomena: A burst of neural activity in the visual cortex may *represent* the face of a friend or neurons in the brain's memory centers may *represent* a childhood memory.

But with the many complex relationships between mind, brain, and environment, it's not always clear what neuroscientists mean when they say neural activity "represents" something. Lack of clarity around this concept can lead to miscommunication, flawed conclusions, and unnecessary disagreements.

To tackle this issue, an interdisciplinary paper takes a philosophical approach to delineating the many aspects of the word "representation" in neuroscience. The work, published in *Trends in Cognitive Sciences*, comes from the lab of Konrad Kording (https://psychology.sas.upenn.edu/people/konrad-kording), a Penn Integrates Knowledge University Professor (https://pikprofessors.upenn.edu/) and senior author on the study whose research lies at the intersection of neuroscience and machine learning.

"The term 'representation' is probably one of the most common words in all of neuroscience," says Kording, who has appointments in the Perelman School of Medicine (http://www.med.upenn.edu/) and School of Engineering and Applied Science (http://www.seas.upenn.edu/). "But it might mean something very different from one professor to another."

It's also a term that is "philosophically loaded," says Ben Baker (https://gsc.upenn.edu/ben-baker), lead study author and a postdoctoral researcher in Kording's lab (http://kordinglab.com/). Discourse on mental representation dates back as far as Aristotle, he says, and those early discussions laid the groundwork for modern cognitive science. However, as neuroscientists adopted the term, its evolving definition was never fully fleshed out.

Using his background in philosophy of mind and philosophy of science, Baker conducted a literature review, analyzing foundational experiments and major themes in neuroscience to determine how researchers applied the term "representation." He and his coauthors compared these usages to the term's function in philosophy and identified three aspects to the use of the word in neuroscience, each of which builds from the previous one.

The first, most straightforward aspect is correlational, in which a neural state correlates with an event or feature of the environment. For example, a researcher may find that a group of neurons fires whenever a mouse sees a predator. Because that particular neural state is *correlated* to seeing the predator, the researcher may say the neural state *represents* the predator.

In the past, technology limited early neuroscientists primarily to recording neural activity, rather than inducing or modeling it. These experiments produced mostly correlative conclusions, which shaped neuroscientists' early concept of representation, according to Kording and Baker.

But those who hope to build comprehensive models of behavior are interested in more than just correlation; they want to know how neural activity *causes* behavior related to a particular event or feature. The second aspect of representation that Baker and Kording identified also includes a causal component. When researchers use the term in this sense, they mean that neural activity related to some event or feature *causes*

behavior related to that event or feature. Building off the previous example, if a mouse's neural state *represents* seeing a predator, then that neural state may *cause* an action such as hiding or running away.

This definition is more common in contemporary neuroscience studies. Modern experimental techniques such as optogenetics allow for precise stimulation of specific populations of neurons, making it easier for scientists to prove causal links between neural activity and behavior.

Beyond *what* behaviors a neural state may cause, many neuroscientists also want to know *why* animals behave that way. The third aspect of representation encompasses this idea by adding a teleological component, which emphasizes an action's purpose rather than its cause.

In this usage, there is a reason the neural state and its corresponding behavior correlate with an event or feature. Returning to the mouse example, the neural activity that *represents* seeing a predator and causes the mouse to run away exists so the mouse will survive. That's its *purpose*.

Teleology is more often discussed in philosophy than in neuroscience, but neuroscientists' reasoning often implies teleology, according to Kording and Baker. Teleological definitions of representation are especially common in neuroscience research that tries to model abstract components of cognitive tasks. Many evolutionary explanations of behavior rely on teleological reasoning as well.

The researchers say they hope their paper will help tamp down ambiguous use of the word "representation" and promote more rigorous discourse in the field. "The end goal is clear communication," Baker says. Once that happens, Kording adds, "there's many different discussions you can have."

Konrad Kording (https://psychology.sas.upenn.edu/people/konrad-kording) is a Penn Integrates Knowledge (https://pikprofessors.upenn.edu/) University Professor with joint appointments in the Department of Neuroscience (http://www.med.upenn.edu/neuroscience/) the Perelman School of Medicine (http://www.med.upenn.edu/) and in the Department of Bioengineering (http://www.be.seas.upenn.edu/) in the School of Engineering and Applied Science (http://www.seas.upenn.edu/).

Ben Baker is a postdoctoral researcher in the Kording lab (http://kordinglab.com/) and a Provost Postdoctoral Fellow (https://research.upenn.edu/postdocs-and-students/penn-provosts-postdoctoral-fellowships/). Baker received his Ph.D. in philosophy from Penn.

Also coauthor on the paper is Benjamin Lansdell (http://kordinglab.com/people/ben_lansdell/index.html), a data scientist in the Department of Developmental Neurobiology at St. Jude Children's Hospital and former postdoctoral researcher in the Kording lab.

Funding for this study came from the National Institutes of Health (awards 1-R01-EB028162-01 and R01EY021579) and the University of Pennsylvania Office of the Vice Provost for Research.

CREDITS
Marilyn Perkins
Writer

DATE
September 30, 2022

SUBTOPICS
Philosophy, Neuroscience, Psychology, Research, Penn Integrates Knowledge Professors

SCHOOLS
Perelman School of Medicine, School of Engineering & Applied Science