

# The Brain from Inside Out: 2020 Kavli Keynote Address Shines Light on Cognition

August 28, 2020



One of the most profound yet least intuitive advances in science over the past century is Albert Einstein's general theory of relativity, which—among other things—defines how space and time are woven together into the fundamental fabric of the cosmos: spacetime. This seemingly esoteric theory has surprisingly direct influence over our everyday lives, from enabling GPS navigation to helping maintain global communication networks.

Remarkably, the idea of spacetime also has an insightful corollary in neuroscience, as explored in APS's 2020 Fred Kavli Keynote Address, presented by György Buzsáki of New York University. In his talk, delivered virtually because of the cancellation of the 2020 APS Annual Convention, Buzsáki challenged the prevailing idea that “time” and “space” are encoded in the brain in separate representations or “coordinates.” Instead, he postulated that neural activity can be described as a succession of events along a space-time continuum.

Buzsáki's central idea is that, rather than having brains that are blank slates at birth, we are born with brains that organize themselves to induce highly structured robust yet flexible patterns. He refers to this as an “inside-out” framework in which the brain comes with a preconfigured and self-organized dynamic that constrains how it acts and how it views the world. In the brain's nonegalitarian

organization, preexisting nonsense brain patterns become meaningful through action-based experience.

To explore this, Buzsáki addressed three interconnected questions:

- How did we inherit our neuroscience framework?
- How did cognitive mechanisms emerge?
- What is the alternative to the current dominating “blank slate” model?

Buzsáki began his talk by revisiting what he proposed is the birth of modern psychology: the 1890 publication of *The Principles of Psychology* by William James. Buzsáki noted that the vocabulary we now use to describe psychology lives within the titles of the chapters in this seminal book. Examples include chapters like “Memory,” “Sensation,” and “Imagination.” These same concepts live today within the established “input-output” model of the brain.

“In this framework, the brain doesn’t change a lot with new experience,” Buzsáki said. “We can learn any amount of novel things without changing our brain dynamic. This is not possible in the tabula rasa, blank-slate model.”

György Buzsáki (New York University)

### **How Did We Inherit Today’s Neuroscience Framework?**

According to the input-output model, the brain learns about the features of the world (inputs), which it then turns into outputs through some sort of intermediary or “black box” function that produces consciousness, decision-making, and free will.

Buzsáki’s main point is to propose an alternative model—the inside-out model—which starts with the idea that the brain is a self-organized system. In this new model, the brain’s main job is to predict the consequences of its action and what is useful for the survival of the body. Starting with existing internally organized patterns, the brain then generates outputs. These outputs, in turn, go on to influence the inputs we receive from the outside world: our perceptions.

In contrast to the input-output model, which goes from specific to general, the inside-out model starts from “good enough” generalizations that become detailed through experience. Buzsáki explained this idea with the simple example of observing a rose.

In the input-output framework (also referred to as the “outside-in” framework), the brain has no grounding for how an object like a rose is related to the rest of the outside world. An inside-out framework, however, provides this needed grounding for cognition by comparing the rose to some action the body “sends out,” like picking up and moving the rose. Whenever the brain sends out an action, it informs the rest of the brain by comparing the changes it perceives in the outside world.

“So my big claim is that this action is the source of grounding, and this is what can give rise to the meaning of many of the perceptions we have,” Buzsáki said.

How is it possible that from a simple brain, we are able to generate a very complex computation that we call cognition?

### **How Did Cognitive Mechanisms Emerge?**

The answer to this question, according to Buzsáki, is that cognition is nothing more than internalized action. In a simple brain model with very few neuronal connections, it is possible to evaluate the environment and predict possible future outcomes with a simple input-output mechanism, but in a very limited manner.

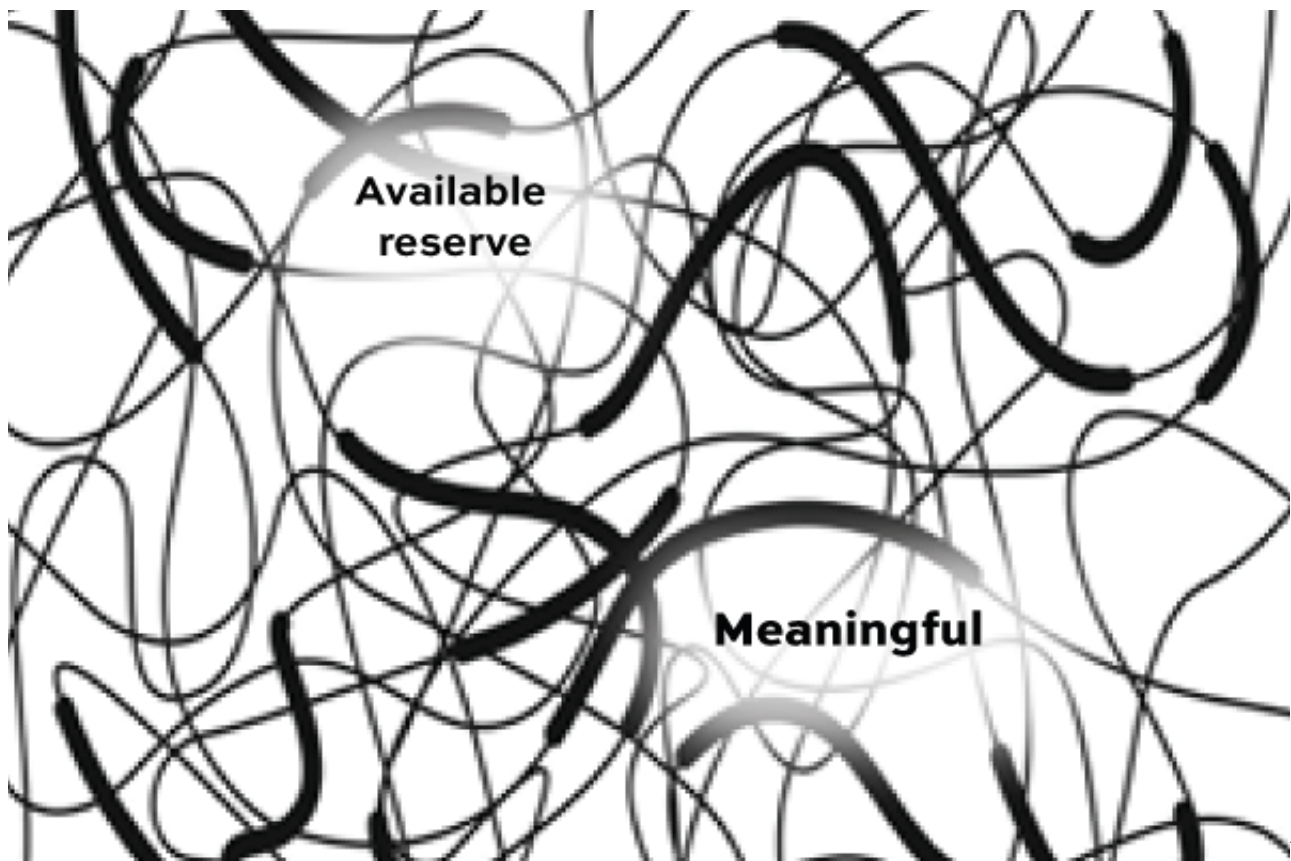
In a more evolved brain, however, the networks become so complex that even without an outside environment, neuronal connections can extrapolate and interpolate. “The brain learns to predict in a much more complex environment and at a much longer time scale,” noted Buzsáki.

Over time, the brain evolves, learns, and becomes capable of no longer needing to interact with the environment to predict outcomes. This enables us to consider “what if” scenarios in which we imagine the consequences of hypothetical actions without acting them out in the real world.

A concrete example of this idea is shown in some of Buzsáki’s earlier work on memory and navigation, which describes two distinct ways of navigating in an environment. The first mode is purely internalized, as when one walks through a completely dark, unfamiliar room. The longer you do this, Buzsáki noted, the more errors accumulate. That is why we need another anchor, which we get when we open our eyes and use landmark, or map-based, navigation.

This same system for navigation is used to create memory in the brain, which we can think of as cognitive navigation. It can also be used to predict the future through imagination and planning. “We are using exactly the same hardware and neuromechanisms as we use in navigation, except we are no longer relying on external landmarks or feedback from the body,” Buzsáki explained.

Just as there are two types of spatial navigation, there are two types of memory that Buzsáki identifies: episodic (self-referenced) and semantic (allocentric). Episodic memory is what, where, and when something happened to someone personally. Semantic memory is when something happens to someone else, or similar events happen many times to the same person but in different contexts. The outside package (specific elements of the experience) is lost, and semantic memory is what is left.



In his talk, Buzsáki discounts the possibility of a blank-slate model, instead proposing that brains come with a preconfigured dynamic and a realm of possible neuronal sequences.

In earlier models of cognition, these foundations of time and space are related to different parts of the brain. In Buzsáki's inside-out model, however, the mental processing of time and space began to merge—a fact that has caused a lot of trouble, as there already existed a definition of episodic memory that establishes the “what” that happened in relation to where it happened and when.

In the outside-in model, recombining these different elements from different parts of the brain is how we create memory. Buzsáki noted that this is similar to Newtonian classical physics, because space is a container and time is an arrow, which gives every experience a time stamp.

This is contrasted, however, with the inside-out model, which has a closer analogy in modern Einsteinian physics. As noted by Carlo Rovelli in 2016: “There is no longer space which ‘contains’ the world, and there is not time ‘in which’ events occur.”

Buzsáki pointed out that neurons firing in a sequence related to the environment are an outside-in mental process. But this type of linear, step-by-step processing doesn't allow for predictive or imaginative inside-out dynamics. “These self-generated, internally generated sequences are the foundations our cognitive ability,” he observed.

To illustrate his point, Buzsáki discussed a study comparing the firing of neurons in an animal's brain before it makes a choice to move left or right toward a reward. The data from this study show that the recorded neuron patterns could predict, with nearly 90% accuracy, the choice that the animal would

make.

The take-home message from this experiment is that every single cell in the hippocampus can function as a place cell or a time cell. Thus, calling neurons in the hippocampus either place cells or time cells is irrelevant to the brain, Buzsáki contended. “What matters is how the downstream reader mechanisms classify hippocampal messages.”

This would mean that episodic memory may need to be redefined, because navigation in both physical and mental space is a succession of events. And the hippocampus is a general-purpose sequence generator that encodes content-limited ordinal structure, suggesting that the hippocampus functions as the brain’s “librarian” or “search engine.”

### **What Is the Alternative to the Blank-Slate Model?**

Buzsáki opened his final point by suggesting that the current model of the brain in neuroscience is akin to a blank-slate model, in which the complexity of the brain “should scale with the amount of experiences you have.” In other words, the brain evolves from very simple to very complex after many years of learning. In this system, the brain starts out with random, egalitarian connections.

The model advocated by Buzsáki, however, is a “skewed,” or nonegalitarian, arrangement of neurons that organize themselves in preformed networks in the brain. With these intrinsic networks, the brain is able to generate an enormous number of sequence patterns without any prior experiences. These connections have a wide range of synaptic weights, firing rates, and population synchrony. Taking his cue from mathematics and the accumulating data, Buzsáki compared the skewed and blank-slate distribution-of-neuron models to logarithmic and linear functions, respectively.

The logarithmic distribution that Buzsáki proposes is the brain’s attempt to reconcile conflicting demands among elements of cognition, including dynamic range, stability, plasticity, and redundancy, among many others.

“They are all competing with each other,” observed Buzsáki, “and nature’s answer to many of these problems is typically diversity.” He demonstrated this idea by showing that, when dealing with a scenario involving many options, like an animal in a seven-arm radial maze, if one neuron fires after selecting one path, the odds are very low that it will fire in another. Some smaller number of neurons, however, fire in every single case.

This suggests that some of these neurons are generalizers, whereas the majority are very specific. But, in an illustration comparing the specific (or downstream) to the general synapses, Buzsáki showed that the generalizers have more extensive connections and fire much faster. So, even though they represent a small fraction of the brain, they can account for about half of the performance in brain processes.

In an experiment that examined learning in animals, Buzsáki and his team found that low-firing neurons are related more to learning; they are plastic and respond readily to specific situations. The minority of fast-firing neurons, however, are more rigid and function as the brain’s “good enough” immediate guess.

This helps us understand how we learn. In the blank-slate model, we start with a simple brain, like blank pages, and we fill in the details. Discarding this possibility, Buzsáki instead concludes that brains come with a preconfigured dynamic and a realm of possible neuronal sequences.

“In this framework, the brain doesn’t change a lot with new experience,” Buzsáki said. “We can learn any amount of novel things without changing our brain dynamic. This is not possible in the tabula rasa, blank-slate model.”

[View the entire APS Fred Kavli Keynote Series here.](#)