

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

## UNDERSTANDING THE CORTEX THROUGH GRID CELLS

---

May-Britt Moser and Edvard I. Moser

One of the ultimate goals of neuroscience is to understand the mammalian cerebral cortex, the outermost sheet of neural tissue that covers the cerebral hemispheres. All mammalian brains have a cortex, but during evolution, the size of the cortex has expanded enormously, and in the largest brains the growth has resulted in extensive folding, with much of the cortical surface getting buried in deep grooves, or sulci and fissures. The cortex is the site where most cognition and intellectual activity takes place. Thinking, planning, reflection, and imagination depend on it. Memories are stored there, and the cortex takes care of language interpretation as well as language production. Moreover, although the cortex can be found across the whole range of mammalian species, the expansion of this brain structure is thought to underlie the amplification of the intellectual repertoire in humans.

### Can We Understand the Cortex?

What is the neural basis of the intellectual functions of the cortex? At its outer limits—such as in deciphering the neural basis of higher brain functions—the cortex may at first seem quite unreachable. But it is important to remember that the cortex also performs more tangible operations, such as interpreting inputs from the sensory environment. The cortical interpretation of sensory signals has for a long time served as the neuroscientist's window into the cortex. By studying, for example, how signals from photoreceptors, olfactory receptors, or touch receptors are represented at early stages of brain processing, in the primary sensory cortices, neuroscientists have made quite significant progress in describing and understanding some of the operational language of the cortex.

A major breakthrough in the analysis of sensory cortices was the discovery of cells that responded selectively to local features of the visual field. In a series of experiments that started at the end of the 1950s, David Hubel and Torsten Wiesel showed that neurons in the primary visual cortex fired specifically when line segments of a specific orientation were presented to the visual field. The linear receptive fields of these cells differed from the circular center-surround fields of cells at earlier stages of processing, in the retina and in the thalamus. Hubel and Wiesel showed that cells with different orientation preferences, or different preferences for inputs from the left and right eyes, were organized in columns of cells with similar functional properties. These findings pointed to a functional architecture for visual computation in the primary visual cortex and provided unprecedented insight into how visual input was fragmented and reassembled at different stages of the visual system and how function was divided across different elements of the visual circuit. Their work started a new era of neuroscience in which the visual cortex served as a guide to cortical computation, with an impact far beyond the direct implications for the mechanisms of vision. Similar progress has subsequently been made in other sensory systems, and we are beginning to understand how senses as diverse as olfaction, taste, and touch are encoded at the level of cortical circuits.

However, while remarkable insights have been made at the early stages of cortical processing, where the first transformations of sensory input take place, little is known about how the brain works at subsequent levels, in the higher-order integrative parts of the cortex—the “high-end cortices.” Yet this is probably the territory of the most challenging cognitive operations, such as the production of thought, decisions, or complex memories. One of the reasons for the inaccessibility of the higher parts of the cortex is that as the distance from the sensory receptors increases, the firing of the neurons becomes increasingly decoupled from the specific features of the sensory environment. As such, it becomes difficult to find correlates in the external world that possess any predictable relationship to the firing pattern of the recorded cells. At the high end of the cortical hierarchy, firing may be triggered via a multitude of converging sensory channels as well as intrinsic processes not corresponding to any particular input. When we do not know the active inputs to a cortical area at any given time, and those inputs originate

from areas whose workings we also do not understand, it is difficult to relate the activity of a cell from one of the high-end cortices to any particular behavioral function.

## The Mammalian Space Circuit: A Window to the High-End Cortices

One exception to the apparent decoupling from the external world is the space-encoding cell population of the hippocampus and the entorhinal cortex, located at the very top of the cortical hierarchy, many synapses away from any of the primary sensory cortices (see color plate 4). In this system, cells have remarkably predictable firing correlates. Many cells in this part of the brain fire only when the animal is in a specific set of locations in its local environment. The preferred locations differ from cell to cell, such that as a population, the cells fire in unique combinations at every location in the environment. Because of these unique activity combinations, the cells effectively serve as a map of the animal's position.

The study of the neural basis of space began in 1971, when John O'Keefe and John Dostrovsky, at University College London, used microelectrodes to record natural activity of neurons in the hippocampus of freely moving rats (figure 1). They were able to pick up impulses, or action potentials, from individual cells in CA1, one of the major sub-fields of the hippocampus. Many of their cells responded specifically to the animal's location in the environment. These cells were named "place cells." When the rat was in the cell's "place field," the cell fired at a high rate. As soon as the rat left this area, the activity decreased and remained low until the next time the animal came to the place field. O'Keefe and his colleagues soon discovered that most hippocampal cells had place fields and that the exact firing locations differed from one cell to the next. Collectively the population of place cells was found to generate a map of the environment, with a unique constellation of active cells at every single position. The strict relationship between neural activity and a property of the environment—the animal's location—was unique among all the recordings that had been made in higher-end cortices by that time.



*Figure 1.* Most of our knowledge of the mammalian space circuit has been obtained from rats and mice. Rodents have a well-developed entorhinal and hippocampal cortex and demonstrate excellent spatial memory and navigation—skills thought to depend on these cortical systems.

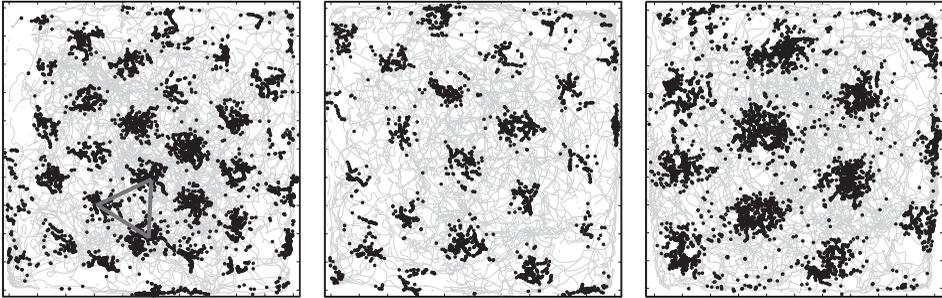
During the decades following the discovery of place cells, accumulating evidence suggested that place cells have functions that extend beyond a specific role in the mapping of the physical space. This idea was reinforced by the observation that different maps could be activated by small variations in the appearance of the environment, suggesting that if the brain has a general map for distances and directions that does not care about what the environment looks like, then such a map should be located elsewhere. Motivated by these considerations, we started, at the turn of the millennium, to search for spatial representations outside the hippocampus. In our first study, with Vegard Brun and a few other graduate students, we recorded from CA1, the hippocampal subregion where place fields had been identified thirty years earlier. In some of the animals we removed the intrinsic connections of the hippocampus, leaving intact only the direct input from the entorhinal cortex, upstream of the hippocampus. Somewhat to our surprise this interference with local hippocampal circuits did not abolish the tendency for CA1 cells to fire in

specific locations. Place cells remained place cells. This implied that unless the place signal was generated entirely by local CA1 processes, the cells must have received critical spatial input from the entorhinal cortex. The findings drew our interests to this unexplored cortical region, one synapse upstream. This brain area turned out to be a gold mine.

## Grid Cells and Grid Maps

In 2004 and 2005, together with our students Marianne Fyhn, Torkel Hafting, and Sturla Molden, and our colleague Menno Witter, we inserted recording electrodes directly in the medial entorhinal cortex, in a part of the area that was strongly connected to the locations in the hippocampus where place cells were normally studied. The findings were quite striking. Individual cells had discrete firing fields, like place cells in the hippocampus, but each cell had multiple fields, and the fields were arranged in a remarkably regular pattern (figure 2). Collectively, the firing fields of an individual cell defined a periodic triangular array covering the entirety of the animal's environment, like the cross points of graphic paper rolled out over the surface of the test arena, but with equilateral triangles as the smallest repeating unit, as on a Chinese checkers board. Because of their grid-like periodic firing pattern of these cells, we named them grid cells. The grid structure was similar for all grid cells, but the spacing of the fields, the orientation of the axes, and the x-y location of the grid fields varied. Grid cells were initially observed in rats, then we found them in mice, and more recently they have been described also in bats, monkeys, and humans, suggesting that they are present widely across the mammalian branch of the phylogenetic tree.

A striking property of the grid cells was the persistence of the firing pattern in the presence of changes in the animal's speed and direction. Moreover, when two grid cells were recorded at the same time, the relationship between their grid fields tended to replicate from one environment to the next. If the grid fields of two simultaneously recorded cells overlapped in one task, they would generally overlap in the next too. The rigidity of this relationship was quite different from the behavior of place cells in the hippocampus, which, based on the work of Bob Muller and John Kubie at SUNY Downstate Medical Center, were known to



*Figure 2.* Grid cells in entorhinal cortex of the rat brain. Three grid cells are shown. Left: cell with short spatial wavelength; right, cell with long spatial wavelength. Each panel shows the trajectory of a foraging rat in a 2.2 m wide square enclosure (gray) with the spike locations of one cell superimposed on the track (black). Each black dot corresponds to one spike. Modified from Stensola et al. (2012).

have the capacity to switch between completely uncorrelated firing patterns. The coherent firing patterns of simultaneously recorded grid cells suggested that the same grid map was used over and over again, pointing to the grid cells—and not the place cells—as a likely implementation of a universal brain metric for space.

However, are all grid cells part of the same map, or are there several maps? In the earliest studies, we recorded from few cells at the time, and those cells were generally from the same location in the entorhinal cortex. It was not possible to infer the structure of the grid map from those limited recordings. In more recent work, with Hanne and Tor Stensola in our lab, we have been able to increase the number of neurons by an order of magnitude. By 2012, we were able to record from more than 180 grid cells across widespread regions of the entorhinal cortex in the same animal. These recordings revealed that grid cells are organized into a small number of maps with discrete properties (see color plate 5). Different grid maps varied on a number of parameters, including the spacing of the grid fields and the orientation of the grid axes. At the dorsal end of the entorhinal cortex, near the top, most grid cells had tightly packed grid fields and all seemed to belong to the same module. As we moved away from the dorsal border, cells from other modules, with a larger grid spacing, joined the ensemble, and at the deepest positions, cells with large grid scales often predominated. A total of four grid

modules were identified in each animal, but the total number may be larger because only a part of the entorhinal cortex was sampled.

What is most striking about the modular organization of the grid cells is that individual modules can respond quite independently to changes in the spatial layout of the environment. Testing animals with cells from four modules in a square box that was compressed to a rectangle (see color plate 6) showed this. Cells from Module 1—the most dorsal module—did not respond to the compression and kept their original firing locations in the common area of the two boxes. In contrast, in cells from larger modules (Modules 2–4), the grid fields were squeezed in one direction in proportion to the shrinkage of the recording box. These observations suggest that, at least in principle, different modules can respond independently of one another when the geometry of the environment is changed. Apparently the grid network consists of four or more discrete maps that may or may not respond in a coherent manner.

What could be the advantage of organizing the brain's map of space in this way? Why would four or more maps be better than just a single coherent map? The answer may lie in the way the grid map is used “downstream” in the hippocampus. While the majority of the cells in the hippocampus are place cells, the hippocampus is also critical for certain types of memory, often referred to as declarative memories. These are memories that we are conscious of—or can “declare”—including memories of facts and events. Space is a fundamental element of these memories. Because we store thousands of declarative memories every day, the hippocampus needs to find a way to keep them all apart. This is where grid modules may be useful. If two modules respond independently to a change in the environment, their coactivity will change. The change in coactivity will activate a new subset of cells in the hippocampus. Each relative displacement among the grid modules may lead to a different activity combination, which in turn may activate a different set of hippocampal neurons. Thus with only a handful of grid modules, it is possible for the entorhinal cortex to link itself to a large number of hippocampal activity patterns, and putative memories, much like the way a combination lock can store 100,000 codes with only 5 counters that each run from 0 to 9. By combining input from a small number of independent grid modules, hippocampal cell populations may acquire the ability to generate huge numbers of discrete representations individualized to specific places and experiences.

## How Are Grid Cells Generated?

One fascinating property of grid cells is that such a regular firing pattern appears so high up in the cortex, far away from the sensory inputs that define the distinct receptive fields of many neurons in the primary sensory cortices. In sensory systems, sensory representations often appear to get more disorderly as the number of synapses from the sensory receptors increases. In contrast, the grid pattern is highly regular, unlike the structure of activity observed so far in areas upstream of these cells. The perfectly hexagonal firing pattern of the grid cells does not correspond to any property of the animal's sensory environment and thus more likely reflects mechanisms that are intrinsic to the entorhinal cortex. What could those mechanisms be—how does a network generate hexagonal firing fields?

While the mechanism of grid formation remains to be established, observations suggest that hexagonal firing patterns emerge as an equilibrium state in competitive networks where all cells inhibit all other cells in their vicinity. Theoretical studies and computational modeling show that in a network where all cells are connected to all other cells within a certain range, via inhibitory connections, hexagonally patterned firing will appear spontaneously as a resting state (see color plate 7). In collaboration with Yasser Roudi and Menno Witter and their colleagues, we have shown that entorhinal cells—in the cell layer that contains the most prototypical grid cells—are connected exclusively via inhibitory interneurons, and that such connections can lead to the formation of hexagonally spaced firing in a model network. Grid cells are perhaps just one of many examples in nature where hexagonal arrangements emerge through self-organizing processes as a result of evenly distributed competitive forces.

## Grid Cells Are Not Alone

Soon after the discovery of grid cells in 2005, it became clear that these cells are not the only spatial cell types in the entorhinal network. Grid cells were the most predominant cell type in the superficial parts of the entorhinal cortex, particularly in the cell layer that contains the strong

inhibitory connections, but with Francesca Sargolini and other students in our lab, we found in 2006 that a large proportion of the entorhinal cell population is direction modulated. These cells, which are similar to a cell type Jim Ranck and Jeff Taube discovered in other brain regions twenty years earlier, fire selectively when the rat points its head in a certain direction. Some of these cells are grid cells at the same time, firing in grid fields only when the animal moves in the cell's preferred direction. In 2008, with Trygve Solstad, we subsequently found that grid cells and head direction cells intermingle with yet another novel cell type—the border cell. These cells fired specifically when the animal was close to one or several borders of the local environment, such as a wall or an edge. When the box was stretched, the firing field followed the wall, and when a new wall was inserted, a new firing field emerged along the insert. Both head direction cells and border cells retained their properties when the animal was moved to a different environment. Two head direction cells that fired in the same direction in one environment tended to fire in the same direction also in other environments, and two border cells with similar wall preferences in one box would have the same preferences also in another box. The rigidity of the head direction and border cells, as well as the grid cells, suggests that the entorhinal maps are used universally across many environments, much unlike the hippocampal place-cell map, which appears to set up new activity combinations for every single environment or experience.

The presence of multiple spatial cell types in the same neural system, such as place cells, grid cells, head direction cells, and border cells, raises some obvious questions. One is how they are related—are place cells formed from grid cells, border cells, or other cells, and are the entorhinal cells, in turn, dependent on place cells? Recent work by Sheng-Jia Zhang and Jing Ye in our lab has shown that the hippocampus receives projections from a variety of entorhinal functional cell types. The most abundant input comes from grid cells, pointing to these as a major source for place information, but also border cells and even cells with no clear spatial correlate project significantly to the hippocampus. How place cells are generated from these inputs remains an open question, but the observations raise the possibility that place cells receive signals from a variety of sources, perhaps in a redundant manner allowing them to respond at specific locations in response to changing sources of inputs. It is also possible that the functional

input to a given place cell varies over time, perhaps with grid cells providing motion-related input at one moment and border cells providing geometric inputs at a different moment. Clear answers to the mechanisms for transformation of signals from one cell type to another will hopefully be obtained during the next few years, considering that experimental tools for addressing such questions are now becoming available.

## What We Can Learn from Grid Cells

With the discovery of place cells and grid cells, as well as other spatial cell types, it has become possible to study neural computation at the high end of the cortical hierarchy, quite independently of sensory inputs and motor outputs. A huge benefit of these cell types is the clear correspondence between the firing pattern and a property of the external world—in this case the animal's location in the environment. The presence of an experimentally controllable firing correlate, combined with the access to multiple discrete cell types, makes it possible to determine not only how each of the firing patterns is generated but also how the firing patterns get transformed from one cell type to the next within the network. Grid cells may not only help us understand how representations are generated in high-end cortices, but such knowledge may also feed back to the sensory cortices, where intrinsic and top-down processes may play a greater role than what was previously appreciated.

The space circuit of the mammalian hippocampus and entorhinal cortex is one of the first nonsensory “cognitive” functions of the cortex that may be understood in mechanistic detail within a not too distant future. Understanding how space is created in this circuit may provide important clues about general principles for cortical computation, extending well beyond the domain of space into the realm of thinking, planning, reflection, and imagination.

## References

- Buzsaki, G., and E. I. Moser. 2013. “Memory, Navigation, and Theta Rhythm in the Hippocampal-Entorhinal System.” *Nature Neuroscience* 16: 130–38.

- Couey, J. J., et al. 2013. "Recurrent Inhibitory Circuitry as a Mechanism for Grid Formation." *Nature Neuroscience* 16: 318–24.
- Felleman, D. J., and D. C. van Essen. 1991. "Distributed Hierarchical Processing in the Primate Cerebral Cortex." *Cerebral Cortex* 1: 1–47.
- Squire, L. R., and S. Zola-Morgan. 1991. "The Medial Temporal Lobe Memory System." *Science* 253: 1380–86.
- Stensola, H., et al. 2012. "The Entorhinal Grid Map Is Discretized." *Nature* 492: 72–78.