Thinking About Brain Cell Assemblies

Howard Eichenbaum

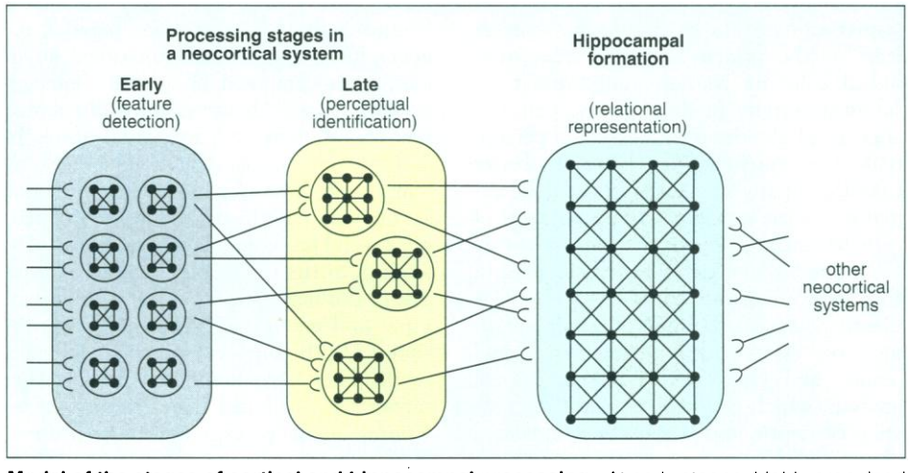
Neurobiologists from Sherrington to Hebb (*1*) have struggled with the problem of how perception, cognition, and memory are coded by assemblies of neurons within the cerebral cortex of the brain. We can now simultaneously monitor the activity of hundreds of neighboring cells (*2*) and have identified functional properties of ensembles of neurons in the cortex and hippocampus (*3–7*). In this issue of *Science*, Wilson and McNaughton (*8*) report one of the first of these studies on the hippocampus, a phylogenetically old part of the brain that is a final stage of information processing for some kinds of memory. They are able to monitor the activity of neural ensembles and show that, when it comes to functional coding, many neurons are better than one. These kinds of studies—plus insights from computational neuroscience—are beginning to show how the hierarchical organization of the cerebral cortex can be coupled with distributed coding. We may soon understand the nature of representation in the "thinking" parts of the brain.

A comprehensive understanding of representation by neural ensembles will require reconciliation of two extreme views of information coding in the cortex. One view espouses a systematic organization composed of a hierarchy of "filters" or "detectors" that encode simple stimulus features and complex events by the activity of single neurons (*9*). A contrasting view espouses a fully distributed representation that encodes each item by distinct spatiotemporal activity patterns of homogeneous arrays of neurons (*10*). These opposing views may in fact be reconcilable; in the brain, both are employed to different extents at successive stages of information processing (see figure).

Evidence in support of the idea that the brain codes information with feature detectors followed closely the development of the extracellular microelectrode in the late 1950s. This technology unexpectedly revealed that cortical firing patterns could readily be related to recognizable dimensions of stimuli and actions (*11*). Many such observations, perhaps most widely promoted by Hubel and Wiesel's studies on the visual cortex, revealed an organized topographical mapping of features and a hierarchy of filtering stages in which the simple features of objects detected at early stages are combined into complex percepts by later processing. By extrapolation, the individual cells at the latest processing stage should represent unique complex objects. This notion has been exemplified by the hypothetical "grandmother cell," a cell activated only by the image of one's grandmother. Indeed, neurons in the inferotemporal cortex, the highest visual area, are maximally responsive to highly specific visual patterns—the image of a monkey's hand (*12*) and, in literal support of grandmother cells, particular faces (*13*). In the hippocampus, the final products of many neocortical processing streams converge. Here, neurons are activated by highly specific and elaborate configurations of stimuli and events, such as the constellation of cues defining the place the animal occupies and the movements made in that place, or the spatial or temporal arrangement of discriminative cues and responses in a training situation (*14*). Thus individual hippocampal neurons can be even more selective than inferotemporal cells; one cell might be most responsive to grandmother in a particular place during a particular family occasion. Taken to this extreme, the organizing principles of specificity, topography, and hierarchy would seem to obviate the need for ensemble coding. However, although this scheme is attractive in its simplicity and organization, it has never been fully accepted, if only because of the obvious requirement for an unrealistically large number of "pontifical" cells to represent all the unique experiences of a lifetime. Serious considerations of these findings have generally taken the hierarchy only to "cardinal" cells that encode higher order dimensions or categorical elements of stimuli, with complex images encoded by the composite activity of a set of such cells (*9*).

Alternatively, scrutinized from the perspective of distributed representation, the data from neurons at each processing stage are also consistent with what one might observe by sampling elements of neural ensembles. Thus, early in processing, each feature is captured by not one but several cells within functionally defined columns of the cortex. Also, most inferotemporal cells do not meet the extreme expectation of specificity; they respond at least somewhat to many visual patterns (*15*). The principle of topography is also compromised in higher cortical areas and replaced by an as yet unclear organization that is reflected in clustering of similar features in large groups of neighboring cells (*13*). In the hippocampus, the principles of specificity and topography are even further modified; hippocampal neurons can have equal specificity but dramatically different selectivity in different behavioral situations in the same environment (*16*) and in the same situation following subtle changes in the environment (*17*).

Topographically, the hippocampus is similar to higher cortical areas in that it represents similar items in clusters of cells, but cells coding for unrelated items overlap,



Model of the stages of cortical and hippocampal processing. At early stages, highly organized sets of small neuronal ensembles (single cortical columns of about 500 μm) detect specific simple features. At late stages, larger ensembles [for example, 1- to 4-mm clusters of face cells in the inferotemporal cortex (*13*)] identify complex images. In the hippocampus, these are represented broadly along with parallel inputs from other cortical systems. The hippocampus acts as a single, very large neural ensemble (15 mm long in primates) that encodes relations among distinct percepts.

The author is at the Center for Behavioral Neuroscience, State University of New York at Stony Brook, Stony Brook, NY 11794.

suggesting an even more complex organizational scheme (*7*). The combination of complex specificities and a breakdown of systematic topography and hierarchy is fully consistent with simulations of neural ensemble performance. Thus, sparse coding of complex features and clustered organization are seen in massively parallel artificial networks (*18*), suggesting that these properties arise even in ensembles that are specifically designed as homogeneous cell assemblies.

Furthermore, real brain cell ensembles appear to utilize population activity in coding processes (*19*). For example, Georgopoulos and his colleagues (*5*) have shown that the population vector computed by amassing the data from several neurons in a single region of primary motor cortex predicts arm movement direction better than data from any single neuron. A simple interpretation of these findings is that, at least in the primary motor cortex, individual cells are coarsely tuned and their summed activity predicts precise performance; similarly, the phenomenon of perceptual hyperacuity (sensory judgment better than that of any single detector) can be explained by the summed activity of groups of cells within modules of the sensory cortex (*20*). Similar approaches have been used in the inferotemporal cortex (*6*). Both a Georgopoulos-like analysis of perceptual dimensions uncovered in a multidimensional scaling and Gerstein's (*21*) gravitational analysis, which does not assume the existence of continuous dimensions, indicated clear face and pattern recognition by population responses that were superior to that by individual cells.

Previous ensemble analyses on the hippocampus have revealed widespread temporal correlations in the activity of the hippocampal cell population (*4*) and nontopographic but clustered organization of functionally similar cells (*7*). Wilson and McNaughton (*8*) have employed a strategy similar to that used in the studies on primate neocortex and have combined trajectory vectors computed from the spatial coding properties of simultaneously recorded hippocampal principal cells. These combined vectors that take into account the activity of about a 100-neuron population are better predictors than information from single cells. Furthermore, the consistency of spatial firing and accuracy of the population prediction increased after initial exploration of a novel environment, while the population code for an adjacent familiar environment was unchanged. These results provide an elegant example of plasticity in hippocampal spatial coding (*17*), and they offer our very first glimpse of the deciphering of a population code in the hippocampus.

Taken together, the findings from empirical and computational research are consistent with an emerging, albeit speculative, view that different combinations of organizational specificity and distributed representation may underlie neuronal ensemble coding at progressive stages of cortical processing (see figure). According to this scheme, early stages of sensory and motor feature analysis involve many parallel, topographically organized, functional modules, each composed of a small homogeneous cell assembly carrying the distributed representation for a particular feature. In contrast, late stages of cortical processing may require a nontopographically arranged set of larger functional modules where neural ensembles combine these features to identify distinct percepts. Finally, as a result of the massive convergence of afferent input and extensive associational connections (*22*), the hippocampus may embody a single, very large functional module that supports a distributed representation of relations among perceptually distinct items (*14*). The findings of Wilson and McNaughton are entirely consistent with this framework, indicating that the same neural space in the hippocampus can encode many items and telling us that hippocampal representation involves the rapid stabilization of ensemble activity. These characteristics of memory representation in the hippocampus might reflect the changing of synaptic weights in the cell assembly in a way meaningful for behavior. Such an ensemble neurophysiological correlate could be used to bridge the gap between the models of synaptic plasticity and real learning and memory (*23*), leading to a comprehensive understanding of brain mechanisms of complex representation.

References and Notes

1. C. S. Sherrington, *Man on his Nature* (Cambridge Univ. Press, Cambridge, United Kingdom, 1941); D. O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949); L. R. Squire, *Memory and Brain* (Oxford Univ. Press, New York, 1987).
2. H. Eichenbaum and M. Kuperstein, *J. Electrophysiol. Tech.* **13**, 189 (1986).
3. A. Michalski, G. L. Gerstein, J. Czarkowska, R. Tarnecki, *Exp. Brain Res.* **51**, 97 (1983); H. J. P. Reitboeck, *IEEE Trans. Biomed. Eng.* **13**, 676 (1983); M. Abeles, G. L. Gerrstein, *J. Neurophysiol.* **60**, 909 (1988); P. M. Gochin, E. K. Miller, C. G. Gross, G. L. Gerstein, *Exp. Brain Res.* **84**, 505 (1991).
4. M. Kuperstein, H. Eichenbaum, T. Van De Mark, *Exp. Brain Res.* **61**, 438 (1986); G. Buzhaki, Z. Horvath, R. Vioste, J. Hetke, K. Wise, *Science* **256**, 1025 (1992).
5. A. P. Georgopoulos, M. Taira, A. Lukashin, *Science* **268**, 47 (1993).
6. M. P. Young and S. Yamane, *ibid.* **256**, 1327 (1992); P. M. Gochin, M. Colombo, G. A. Dorfman, G. L. Gerstein, C. G. Gross, *J. Neurophysiol.*, in press.
7. H. Eichenbaum, S. I. Wiener, M. Shapiro, N. J. Cohen, *J. Neurosci.* **9**, 2764 (1989).
8. M. A. Wilson and B. L. McNaughton, *Science* **261**, 1055 (1993).
9. H. B. Barlow, *Q. J. Exp. Psychol.* **A 37**, 121 (1985).
10. W. J. Freeman, *Mass Action in the Nervous System* (Academic Press, New York, 1975).
11. V. B. Mountcastle, *J. Neurophysiol.* **20**, 408 (1957); D. H. Hubel and T. N. Wiesel, *J. Physiol. (London)* **160**, 106 (1962).
12. C. G. Gross, C. E. Roche-Miranda, D. B. Bender, *J. Neurophysiol.* **35**, 96 (1972).
13. D. I. Perret, A. J. Mistlin, A. J. Chitty, *Trends Neurosci.* **10**, 358 (1987).
14. H. Eichenbaum, T. Otto, N. J. Cohen, *Behav. Neural Biol.* **57**, 2 (1992).
15. R. Desimone, T. D. Albright, C. G. Gross, C. Bruce, *J. Neurosci.* **4**, 2051 (1984); B. J. Richmond and L. Optican, *J. Neurophysiol.* **57**, 147 (1987).
16. S. I. Wiener, C. A. Paul, H. Eichenbaum, *J. Neurosci.* **9**, 2737 (1989).
17. E. Bostock, R. U. Muller, J. L. Kubie, *Hippocampus* **1**, 193 (1991); K. B. Austin, W. F. Fortin, M. L. Shapiro, *Soc. Neurosci. Abstr.* **16**, 263 (1990).
18. M. L. Shapiro and P. A. Hetherington, *Behav. Neurosci.* **107**, 34 (1993); D. Zipser and R. A. Andersen, *Nature* **331**, 679 (1988); R. Linsker, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8390 (1986).
19. In some of these studies, ensemble activity was collected by collapsing the data from one or a few cells recorded at a time while the animal performs a highly repetitive task. This technique is likely valid for experiments that require only a coarse temporal resolution.
20. P. S. Churchland and T. J. Sejnowski, *The Computational Brain* (MIT Press, Cambridge MA, 1992).
21. G. L. Gerstein, D. H. Perkel, J. Dayhoff, *J. Neurosci.* **5**, 881 (1985).
22. D. G. Amaral and M. P. Witter, *Neuroscience* **31**, 571 (1989).
23. H. Eichenbaum and T. Otto, *Trends Neurosci.* **16**, 163 (1993).