Coding for visual categories in the human brain

Charles G. Gross

Single-neuron recordings in the human hippocampus, entorhinal cortex and amygdala demonstrate that cells in these areas can respond selectively to particular categories of visual stimuli.

For the last 50 years, the microelectrode has been one of the most powerful tools for revealing the neuronal bases of perception and cognition. Electrophysiology has identified individual neurons that respond selectively to highly complex and abstract visual stimuli. The first discovered and most intensively studied are the 'face-selective' neurons in inferior temporal (IT) cortex of the macaque monkey[1]. Face-selective cells, which are a small proportion of IT cells, respond best or only to the sight of faces. Although virtually all faces activate these cells, different faces produce varying response magnitudes. Thus, the response of one such cell cannot represent a specific face, but the pattern of firing across a set of face-selective cells would be unique for a particular face. As these cells usually do not code for individual faces, they are best described as 'face-category' cells. The brain may represent individual faces by the pattern of activity over a population of such face-category cells, in what is termed population or coarse coding[2,3].

Almost all fundamental questions in neuroscience derive from human experience, and yet, for ethical and practical reasons, they can usually be investigated only in nonhuman animals. Thus, despite our Darwinian faith in the continuity of species, it is always a relief to confirm in humans a basic finding from other animals. In a bold and imaginative study in this issue, Kreiman and colleagues[4] confirm and extend results from monkeys. They report single neurons selective for faces and other visual categories in the medial temporal lobe of humans. This study is interesting and important for two reasons. First, the authors found neurons selective for a variety of visual categories in addition to faces. Second, these neurons were found in areas of the temporal lobe, which unlike IT cortex are not known to be specialized for representing visual stimuli or visual categories, and are more commonly associated with memory.

The subjects were patients with epilepsy who were resistant to drug treatment. To find an epileptogenic focus that could be surgically excised, the surgeons implanted patients with intracranial depth electrodes for several weeks. The electrodes were in portions of the medial temporal lobe, including the hippocampus, entorhinal cortex and amygdala. (A possible problem with such studies is that the recordings are made from diseased brains, and indeed quite near known sites of malfunction.) While action potentials were recorded from single neurons, patients viewed complex visual stimuli that fell into one of nine categories: drawings or photographs of famous people, emotional expressions on unknown actors, household objects, cars, animals, food, spatial layout (scenes) and abstract patterns (Fig. 1). Several examples from each category were shown several times during the recording session with each neuron. To insure that the stimuli engaged the patients' attention, they had to indicate whether the picture was a face or not by pressing a button. This type of experiment is more restricted in humans than in monkeys because the electrodes cannot be moved around to sample different neurons, and the time available to study the activity of each is severely limited.

Kreiman and colleagues[4] found that 18% of hippocampal, 16% of entorhinal and 9% of amygdala neurons responded selectively to one or, more rarely, to two of the visual categories. For example, some neurons fired more in response to pictures of animals than to stimuli in any other category. However, no particular animal elicited more firing than any other animal. Other neurons fired more in response to both photographs and drawings of famous faces but not to any other categories; again, there were no differences in responses to different face stimuli. Thus, these neurons

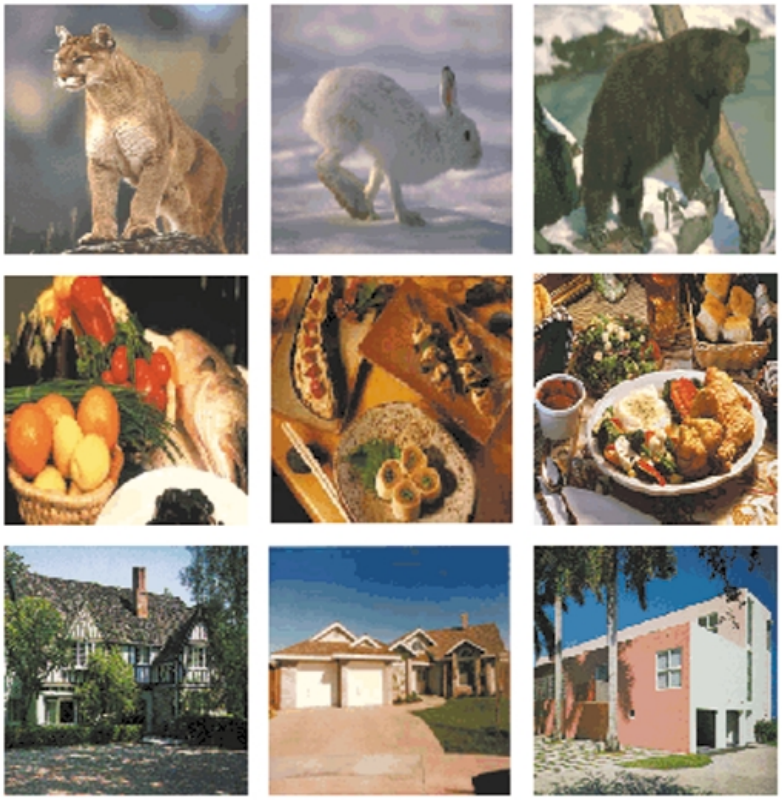


Fig. 1. Examples from three of the nine categories of visual stimuli used by Kreiman and colleagues[4]. Top row, animals; middle row, food; bottom row, spatial layout (scenes).

The author is at the Department of Psychology, Princeton University, Green Hall, Princeton, New Jersey 08544, USA.
e-mail: cggross@princeton.edu

may be described as 'category selective'. Neurons in human hippocampus and amygdala are known to be selective for faces, objects or letters[5,6], but the current results[4] more extensively demonstrate selectivity for visual categories.

The existence of category-selective cells fits nicely with other studies of human brain function. After brain damage, particularly to the temporal lobe, humans may have selective deficits in visually recognizing specific categories of objects. For example, in spite of otherwise normal perception, patients may have selective difficulty in recognizing only faces, only animals or only man-made objects[7]. Similarly, PET and fMRI reveal localized regions of the temporal lobe that are differentially active in response to specific categories of visual stimuli such as faces, words, houses or chairs[8].

Kreiman and colleagues found that hippocampal cells were more selective for spatial scenes than for any other visual categories, which was not the case for either amygdala or entorhinal neurons. This is consistent with long-standing evidence for specialization of the rodent hippocampus for spatial processing[9], which seems to be true in humans as well[10].

In monkeys, neurons in IT cortex are selective not only for faces, but also for other natural categories of visual stimuli, such as hands[11] and facial expressions[12]. Other studies show selectivity for arbitrary visual categories after explicit training to distinguish those categories. For example, Logothetis and colleagues showed that monkeys trained to discriminate wire figures had IT neurons selective for such figures but not for spheroidal objects, and the converse was true for monkeys trained on spheroidal objects[13]. Thus neurons selective for visual categories are found in both humans and monkeys, for natural and arbitrary categories. Although most of these categories are presumably learned, some, such as selectivity for faces and facial expression, may well be present at birth in both humans and monkeys[14,15].

One difference between human and monkey studies of category-selective neurons is in the recording sites. Virtually all systematic analyses of stimulus selectivity in temporal lobe neurons in the monkey has been on IT neurons rather than on hippocampal, entorhinal or amygdala cells. Perhaps this is because IT cortex is assumed to be the last exclusively visual processing station in the ventral cortical pathway, which analyzes information about visual identity. IT cortex is believed to send the result of its analysis to the hippocampus (by way of entorhinal cortex), and the hippocampus is thought to be involved in consolidation of short-term memories into long-term ones. Thus, monkey studies on hippocampal and entorhinal cells have almost exclusively been concerned with questions of supra-modal recognition memory, short-term memory and similar mnemonic matters, rather than selectivity to visual stimuli and their categorization.

What are the implications of Kreiman and colleagues' demonstration[4] of category-selective cells in medial temporal areas? One is that the hippocampus carries more than just relational or spatial information (although the hippocampus had the largest proportion of cells selective for the category of 'spatial scenes'). A related implication is that the hippocampus has more than just a modulation or consolidation effect on cortex; instead, it carries complex visual information.

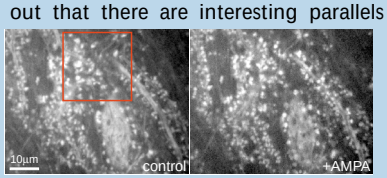
The evidence that medial temporal cells are selective for visual 'categories' implies that these cells may be involved in visual categorization. In addition, and perhaps more parsimoniously, the category-selective cells may be members of an ensemble that represents individual category members by population or sparse coding. For example, the cross-fiber pattern of firing of the ensemble of cells selective for animals may be the code for individual animals, just as face cells seem to form ensembles for encoding individual faces.

The authors' conclusion that their results "may be relevant in the representation and retrieval of visual information" is an appropriately modest one. However, describing the stimulus or even the category selectivity of temporal neurons is still very far from understanding how the circuits they compose underlie the complexity and subtlety of perception and the mysteries and vagaries of memory.

1. Gross, C. G. *Cereb. Cortex* 5, 455–469 (1994).
2. Desimone, R. *J. Cogn. Neurosci.* 3, 1–8 (1991).
3. Gross, C. G. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 335, 3–10 (1992).
4. Kreiman, G., Koch, C. & Fried, I. *Nat. Neurosci.* 3, 946–953 (2000).
5. Heit, G., Smith, M. E. & Halgren, E. *Nature* 333, 773–775 (1988).
6. Fried, I., MacDonald, K. A. & Wilson, C. L. *Neuron* 18, 753–765 (1997).
7. Warrington, E. K. & McCarthy, R. A. *Brain* 110, 1273–1296 (1987).
8. Martin, A., Ungerleider, L. G. & Haxby, J. V. in *The New Cognitive Neurosciences* (ed. Gazzaniga, M. S.) 1023–1036 (MIT Press, Cambridge, Massachusetts, 2000).
9. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Oxford Univ. Press, 1978).
10. Maguire, E. A., Frackowiak, R. S. J. & Frith, C. D. *J. Neurosci.* 17, 7103–7110 (1997).
11. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. *J. Neurophysiol.* 35, 96–111 (1972).
12. Hasselmo, M. E., Rolls, E. T. & Baylis, G. C. *Behav. Brain Res.* 32, 203–218 (1989).
13. Logothetis, N. K. & Sheinberg, D. L. *Annu. Rev. Neurosci.* 19, 577–621 (1996).
14. Fantz, R. in *Behavior of Non-Human Primates* Vol. II (eds. Schrier, A. M., Harlow, H. F. & Stollnitz, F.) 365–403 (Academic, New York, 1965).
15. Rodman, H. R., Scalaidhe, S. P. & Gross. C. G. *J. Neurophysiol.* 70, 1115–1136 (1993).

The mechanics behind spines on the move

Dendritic spines—the main sites of excitatory synaptic contacts in the CNS—have been in the spotlight recently. Their curious motility has been linked to synapse formation as well as plasticity in response to sensory experience. Now, Matus and colleagues (pages 887–894, this issue) have directly addressed the mechanism by which neural activity and spine motility may be linked. The authors made time-lapse videos of spine motility in GFP-actin transfected hippocampal neurons as well as slice cultures from transgenic mice expressing GFP-tagged actin. A major finding was that actin dynamics are rapidly and reversibly inhibited following activation of AMPA receptors—spines became more stable and assumed a more regular appearance. Furthermore, inhibition of motility via AMPA receptors required postsynaptic membrane depolarization and the influx of calcium. In combination with previous work, the results suggest that spines initially formed by NMDA receptor activation are subsequently stabilized by AMPA receptors. Although the results may seem paradoxical because the quite different processes of spine formation and stabilization both require the influx of calcium, the authors point out that there are interesting parallels with growth cone motility, where calcium activation at different stages of synapse formation can have opposite effects.



John E. Spiro