Seeking Categories in the Brain

Simon J. Thorpe and Michèle Fabre-Thorpe

Perceptual categorization is a fascinating cognitive operation in which the mammalian brain groups together objects that share common properties, regardless of their physical differences. For example, we naturally group together cats, fish, birds, insects, and snakes into the category "animal," even though visually they are very diverse. Understanding categorization is a major challenge facing cognitive neuroscientists, a challenge that Freedman and co-workers (1) take on in their study on page 312 of this issue.

These authors examined the responses of neurons in the prefrontal cortex (PFC) of monkeys trained to categorize animal forms (generated by computer) as either "doglike" or "catlike." By continuously "morphing" the basic form of one animal into the other, the authors were able to test (with single-cell recording electrodes) how monkey PFC neurons responded to forms that could be either cat or dog (that is, shapes that were somewhere between the two animals). They report that many PFC neurons responded selectively to the different types of visual stimuli belonging to either the cat or the dog category and with the same strength, regardless of how morphologically close the images were to the other category. The firing of impulses by PFC neurons thus reflects category membership rather than simple processing of the physical characteristics of the images.

The neurons that Freedman et al. recorded from almost certainly receive their visual inputs from the inferior temporal cortex (ITC), a part of the brain that lies at the end of the chain of visual processing stages of the so-called ventral visual pathway (see the figure). It has been known for many years that some ITC cells can be highly selective to particular visual stimuli such as faces (2, 3) and can even respond to a range of two-dimensional views of the same object (4). More recently, Vogels examined the responses of ITC cells in monkeys trained to categorize pictures of trees and fish. He reported a number of cells that were only activated by certain stimuli belonging to a given category (5), although none of them responded to all exemplars of the category. In a particularly impressive recent study, Sheinberg and Logothetis recorded the activity of ITC neurons in monkeys trained to search a large color photograph for small hidden figures—very much like the "Where's Waldo" game familiar to children (6). A wide range of different objects were artificially divided into two sets. To get a reward, the monkey had to pull a lever on the left for one set and on the right for the other set. The monkeys were extremely good at the task, and many ITC neurons showed a strong burst of firing when the monkey's eyes landed on (or close to) particular targets, remaining silent while the monkey was exploring the rest of the natural scene. However, there was no obvious relation between the set of targets to which the neuron responded and the artificial object categories as defined by the two response sets. It thus appears that the cognitive task of the ITC cells may be different from that of the PFC neurons described by Freedman et al.—activity patterns in the Freedman monkey neurons changed when the same set of images needed to be categorized in a different way. Clearly we need experiments that directly compare ITC and PFC responses using the same behavioral tests. Nevertheless, it looks like ITC and PFC may have different parts to play in these higher order visual tasks: ITC may provide highly processed visual information concerning the visual objects that are present, but PFC may be required to decide how these objects should be categorized.

In a way, this distinction between the visual representations seen in ITC and the more behaviorally relevant activity in PFC

The authors are at the Centre de Recherche Cerveau et Cognition UMR 5549, Université Paul Sabatier, Toulouse, 31062 France. E-mail: thorpe@cerco.ups-tlse.fr

is reminiscent of some much earlier work on visual responses to food. These studies showed that neurons in the lateral hypothalamus (an area in the limbic system involved in the control of feeding) can respond to all stimuli that the monkey treats as food [7]. In contrast, neurons in ITC, which probably provide the input for these food-selective neurons, fail to show such category-relatedness [8] despite their considerable stimulus specificity.

One of the most impressive features of visual responses seen in both PFC and ITC is their speed. In ITC, neurons start to respond about 100 ms after stimulus onset, and in PFC typical onset latencies are only slightly longer. Although 100 ms may seem like a fair amount of time, it is not very long when one takes into account the number of processing stages involved (see the figure). Information from the retina reaches the primary visual cortex, area V1, via the thalamus, and is subject to further processing in areas V2 and V4 before reaching the various parts of ITC and then PFC. Response properties become more and more complex as one moves along this ventral cortical stream, and onset latency increases in a fairly systematic way, with an increase of roughly 10 ms per stage. This does not allow much time for complex iterative processing and suggests that the initial activation of cells in ITC and PFC could depend largely on a feedforward pass through the visual system. Oram and Perrett provided support for such a view by showing that even the earliest part of the response of ITC neurons could be highly selective [9]. But even stronger evidence for a feedforward mechanism comes from another recent study that examined the response of ITC neurons to strings of images presented in rapid succession (a technique borrowed from experimental psychology known as RSVP, rapid serial visual presentation) [10]. Even when the images were changing at 72 Hz (a new image every 14 ms), ITC neurons were still able to follow the input through a statistically significant modulation of their discharge each time their preferred stimulus was shown. This kind of data has strong implications for our understanding of visual processing because it implies that the visual pathway must be acting as a sort of pipeline processor, with different images being processed simultaneously at different levels of the system.

On the other hand, other recent findings suggest that not all visual information can be analyzed on the basis of this first wave of information processing. Although the initial response of ITC neurons is capable of signaling whether a face is present, other types of information, such as

facial expression or identity, are only available later on [11]. However, in some categorization tasks, the behavioral reaction times can be so short that the decision is presumably taken without waiting for this later process to conclude. Freedman et al. report a mean reaction time of 264 ms, which matches the values seen with a go/no-go animal categorization task using briefly flashed photographs [12, 13]. But, when the mean reaction time is 250 to 260 ms, some responses can be reliably pro-

times of human volunteers to be reduced even further, to around 230 to 250 ms [15], this value is still roughly 50 ms longer than monkey reaction times.

A number of electrophysiological studies have shown category-specific activity in humans, but activity appears to be somewhat later than in monkey ITC and PFC neurons. Differential brain activity between target and nontarget trials has been reported in human volunteers about 150 ms after stimulus presentation in a va-
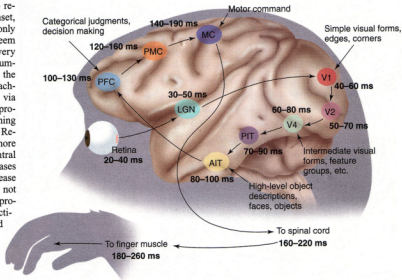
From input to output. Monkeys can categorize complex visual stimuli very quickly, with reaction times that average 250 to 260 ms but that can be as short as 180 ms. Depicted is a plausible route between the retina and the muscles of the hand during a categorization task. Information from the retina is relayed by the lateral geniculate nucleus (LGN) before reaching V1, the primary visual cortex. From there, processing continues in areas V2 and V4 of the ventral visual pathway before reaching visual areas in the posterior and anterior inferior temporal cortex (PIT and AIT), which contain neurons that respond specifically to certain objects. The inferior temporal cortex projects to a variety of areas, including the prefrontal cortex (PFC), which contains the visually responsive neurons that categorize objects [1]. To reach the muscles in the hand, signals probably need to pass via the premotor cortex (PMC) and primary motor cortex (MC) before reaching the motor neurons of the spinal cord. For each processing stage, two numbers (in milliseconds) are given: The first is an estimate of the latency of the earliest neuronal responses to a flashed stimulus, whereas the second provides a more typical average latency.

duced with reaction times as short as 180 ms. This is particularly impressive given that it is only 80 ms longer than the onset latency of typical ITC neurons. For these go/no-go categorization tasks essentially the same task that we have used to determine the speed of visual processing in humans [14]. Interestingly, monkeys appear to be able to produce behavioral responses that are substantially faster than those of even the fastest humans. Although recent replications of the basic scene categorization task have allowed the shortest reaction

riety of categorization tasks [16]. These include recognizing nonanimal [14, 17, 18] face versus nonface [19, 20], and even means of transport [15]. Here again, there are strong grounds for believing that such category-related activity results from feedforward processing. One such argument comes from the fact that neither the onset latency of animal-specific differential activity nor the latency of the shortest reaction times are any faster for very familiar images versus images that have never been seen before [21]. Thus, it appears that even

extensive contextual information is unable to increase processing speed—perhaps because neuronal processing is already so optimized that there is no room for further improvement.

Could the category-specific activation reported in monkey ITC and PFC at around 100 ms correspond to the 150-ms activation seen in humans? Similarities between the monkey and human brain are difficult to establish, but those between monkey ITC and the more ventrally located human fusiform gyrus (where much of the category-related activation seems to be generated) are striking. Why is it, then, that the onset latencies differ between the two species? One possible reason is simply that the monkey brain is smaller than ours. There is not a great deal of detailed information available, but the conduction velocity of intracortical axons used to send information from V1 to V2 to V4 to ITC could be relatively slow, perhaps only 1 to 2 m/s (22). This means that quite a lot of time may be taken up by simply getting information from A to B—a problem that is less serious when your brain is smaller.

But the question still remains whether the category-specific activity seen in humans corresponds to categorization of the type described by Freedman et al. in monkeys, in which the boundaries between categories are specifically coded by single cells. The alternative is that the strong responses recorded from structures such as the fusiform gyrus in humans reflect the activity of large overlapping populations of neurons tuned to particular sets of objects, as appears to be the case in monkey ITC. The most direct test requires single-cell recording from individual neurons. Although normally this is not possible in humans, intracerebral recording in patients with severe epilepsy recently allowed progress to be made. For example, recording of individual neurons in the human medial temporal lobe revealed neuronal responses that were selective not only for faces, but also for natural scenes, houses, famous people, and animals (23).

These new data—regardless of whether they represent the rapid selective visual responses of ITC and PFC neurons in monkeys, the rapid category-specific signals seen in humans, or the fast behavioral reaction times seen in both species—pose a major problem for current models of visual processing. In particular, they imply that a great deal of processing can be done on the basis of a largely automatic feedforward pass through the visual system. In a sense, the fact that visual categorization is fast and robust is perhaps not so surprising. We all have the impression that as we zap from channel to channel, the moment when we categorize what the image contains is virtually instantaneous. The problem now is to understand how the brain can perform this task so quickly and efficiently with neurons that fire electrical impulses 10 million times less rapidly than the transistors in today's desktop computers.

References

1. D. J. Freedman, M. Riesenhuber, T. Poggio, E. K. Miller, *Science* **291**, 312 (2001).
2. C. Bruce et al., *J. Neurophysiol.* **46**, 369 (1981).
3. D. I. Perrett et al., *Exp. Brain Res.* **47**, 329 (1982).
4. M. C. A. Booth, E. T. Rolls, *Cereb. Cortex* **8**, 510 (1998).
5. R. Vogels, *Eur. J. Neurosci.* **11**, 1239 (1999).
6. D. L. Sheinberg, N. K. Logothetis, *J. Neurosci.*, in press.
7. E. T. Rolls et al., *Brain Res.* **130**, 229 (1977).
8. E. T. Rolls et al., *Brain Res.* **164**, 121 (1979).
9. M. W. Oram, D. I. Perrett, *J. Neurophysiol.* **68**, 70 (1992).
10. C. Keysers et al., *J. Cogn. Neurosci.*, in press
11. Y. Sugase et al., *Nature* **400**, 869 (1999).
12. A. G. Delorme et al., *Vision Res.* **40**, 2187 (2000).
13. M. Fabre-Thorpe et al., *Neuroreport* **9**, 303 (1998).
14. S. Thorpe et al., *Nature* **381**, 520 (1996).
15. R. Van Rullen, S. J. Thorpe, *J. Cogn. Neurosci.*, in press.
16. A. M. Treisman, N. G. Kanwisher, *Curr. Opin. Neurobiol.* **8**, 218 (1998).
17. A. Antal et al., *Brain Res. Cogn. Brain Res.* **9**, 117 (2000).
18. J. S. Johnson et al., *Soc. Neurosci. Abstr.* **26**, 952 (2000).
19. T. Allison et al., *Cereb. Cortex* **9**, 415 (1999).
20. D. A. Jeffreys, *Vis. Cognit.* **3**, 1 (1996).
21. M. Fabre-Thorpe et al., *J. Cogn. Neurosci.*, in press.
22. L. G. Nowak, J. Bullier, in *Extrastriate Cortex in Primates*, J. Kaas, K. Rockland, A. Peters, Eds. (Plenum, New York, 1997), pp. 205–241.
23. G. Kreiman et al., *Nature Neurosci.* **3**, 946 (2000).