NeuroReport 7, 2757–2760 (1996)

Rapid visual learning in neurones of the primate temporal visual cortex

Martin J. Tovee[1], Edmund T. Rolls[CA] and Vilayanur S. Ramachandran[2]

University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD; [1]Department of Psychology, University of Newcastle, Ridley Building, Newcastle upon Tyne, NE1 7RU, UK; [2]Center for Research on Brain Cognition, 0109 University of California San Diego, La Jolla, CA 92093, USA

[CA]Corresponding Author

THE human visual system can learn to recognize visual stimuli rapidly. For example, humans can accurately reconstruct meaningful objects out of fragmentary evidence, once they have seen the same object in its unambiguous form. The anterior temporal cortical areas of macaques contain some neurones with invariant visual responses which appear to provide a representation of complex patterns and objects, such as faces. Remarkably, these neurones show an enhancement of response after brief (e.g. 5 s) exposure to the unambiguous stimulus, an effect that appears to reflect the neural basis of the rapid perceptual learning seen in humans.

Key Words: Inferior temporal visual cortex; Learning; Plasticity; Recognition; Vision

Introduction

The human visual system can learn to recognize visual stimuli rapidly. For example, humans can accurately reconstruct meaningful objects out of fragmentary evidence[1] (A.Treisman, personal communication; Fig. 1), but naive subjects often experience a long latency before the object is recognized unless they have been previously exposed to an unambiguous version of exactly the same object embedded in the appropriate context. Once they have recognized the unambiguous version, however, even the fragments evoke the perception of the object virtually instantly, a remarkable example of perceptual learning. The inferior temporal cortex and related visual areas in the cortex in the anterior part of the superior temporal sulcus contain neurones with invariant visual responses which appear to provide a representation of complex patterns and objects, such as faces.[2–5] The issue arises of how this representation is organized. In the experiments described here, we investigated how this rapid form of perceptual learning affects the responses of neurones in this part of the temporal lobe visual cortex. The learning paradigm used ambiguous stimuli that are difficult to recognize until an unambiguous version of the same stimulus is shown. Examples of the unambiguous and ambiguous images used are shown in Figure 1.

Materials and Methods

Twenty-one experiments were performed on 21 single visual neurones that were preferentially responsive to images of faces, recorded with conventional techniques described elsewhere[6] in the cortex in the banks of the anterior part of the superior temporal sulcus and in the inferior temporal cortex (typical sites are shown in Refs 6–8) in two macaques (*Macaca mulatta*) performing a visual fixation task. Correct fixation was confirmed using the search coil technique. The criteria for classifying neurones as face selective have been detailed elsewhere,[6,9] and the 21 neurones on which the experiments were completed were found among several hundred neurones recorded. Examples of two of the grey scale images, and of their binarized ambiguous versions, are shown in Figure 1.

The neurones initially received 10 0.5 s presentations of an ambiguous image of a face which had never been seen before, interleaved with 10 presentations of a control image (trial block 1). This was followed by 10 presentations of the same face but in its unambiguous (grey scale) form interleaved with control images (trial block 2). There were then 10 presentations of the original ambiguous face used in trial block 1 and the control images (trial block 3). Further trial blocks were run if necessary. The control images were either ambiguous and unambiguous image pairs of complex

FIG. 1. Examples of the unambiguous and ambiguous images of faces and complex objects used in the experiment. The stimuli used were grey scale images shown on a video monitor which subtended 8.5° of visual field. Each presentation of an image lasted for 0.5 s, and the images to be shown were chosen in a random sequence. The unambiguous images were grey scale pictures of faces or of complex objects taken against a patterned background. The ambiguous images were binarized black/white versions of the unambiguous pictures which tended to camouflage the faces or objects so that they can no longer be segmented easily from the background. These images resembled Mooney face images.

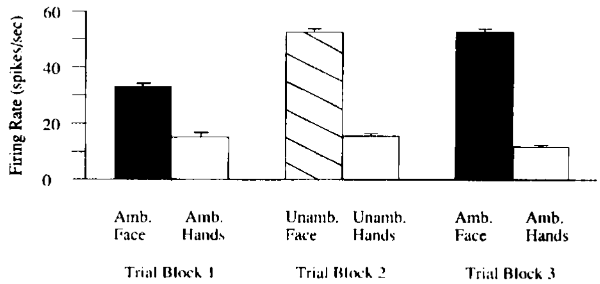

FIG. 2. For the first set of 10 presentations of each stimulus, the ambiguous face elicited firing of 33 spikes s^{-1} (averaged in all cases over the 10 trials in a block), and the control image, which was an ambiguous image of hands, elicited firing of 15 spikes s^{-1}. In stimulus block 2, the unambiguous (i.e. grey scale) version of the same face elicited firing of 53 spikes s^{-1} and the unambiguous version of the control stimulus (hands) elicited firing of 16 spikes s^{-1}. In trial block three, the ambiguous version of the face elicited a much greater firing rate (53 spikes s^{-1}), whereas the firing to the control ambiguous version of the hands did not increase (13 spikes s^{-1}). A one-way analysis of variance and *post hoc* Newman–Keuls' tests showed that in trial block 3 there was a significant increase in the response to the ambiguous face compared with the response to the same image in trial block 1, and that there was no similar increase to the ambiguous version of the hands. A two-way analysis of variance where one factor was face *vs* control image, and the other was firing in stimulus block 1 *vs* firing in stimulus block 3, showed that there was a significant interaction ($F[34,1] = 14.8$, $p << 0.001$). The experiment thus indicates that seeing the face in its unambiguous form (in trial block 2) increased the response of the neurone to the ambiguous version of the same face (seen later in trial block 3). The unchanged response to the control image shows that there was not simply an increased responsiveness to all images produced by showing the unambiguous images in trial block 2.

objects, or ambiguous face images presented without being preceded by their unambiguous counterparts.

Results

A significant increase in firing to the cued ambiguous images was measured in seven of 21 independent experiments, examples of which are shown in Figures 2–4. For the cell shown in Figure 2, the experiment indicates that seeing the face in its unambiguous form (in trial block 2) increased the response of the neurone to the ambiguous version of the same face (seen later in trial block 3). The unchanged response to the control image shows that there was not simply an increased responsiveness to all images produced by showing the unambiguous images in trial block 2. The data from single trials of the same experiment are shown in Figure 3.

The experiment shown in Figure 4 indicates that the neurone increased its response in trial block 3 only to the face which had been shown in unambiguous form in trial block 2.

In the seven cases in which learning was demonstrated the interactions were usually highly significant ($p < 0.001$, 0.0026, 0.0023, 0.001, 0.00001, 0.039, 0.0006 respectively). The chance of this set of seven results being this significant in 21 experiments was

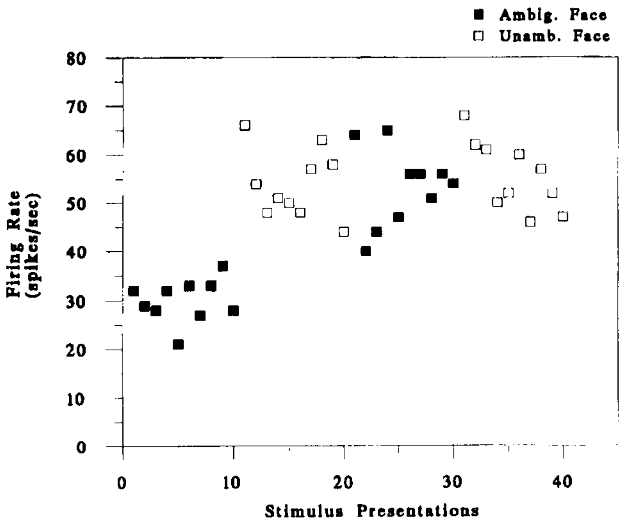

FIG. 3. The responses of the same cell as in Fig. 2, to the three sets of stimulus presentation. Each square represents the response of the neurone to a single stimulus presentation. The empty squares represent the responses to the ambiguous face and the filled squares the responses to the unambiguous version of the same face. The change in firing did not represent a gradual drift upwards in the strength of response over the course of the second 10 presentations, but a sharp and significant change starting at the first response.

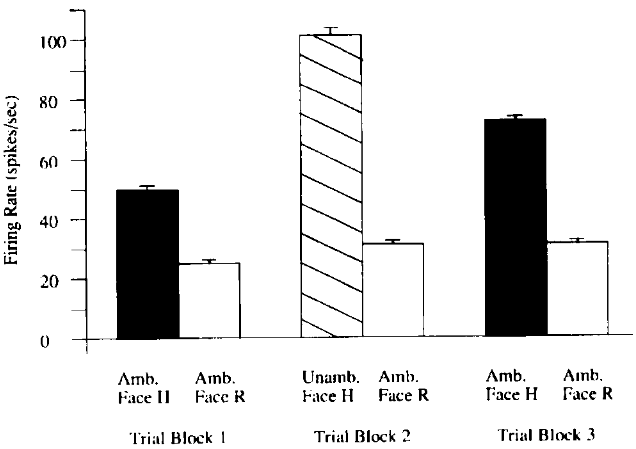

FIG. 4. In trial block 1, the neurone responded to ambiguous face H at 50 spikes s^{-1}, and to control ambiguous face R at 25 spikes s^{-1}. In trial blocks 2 and 4, the unambiguous version of face H was used, but the control image remained as the ambiguous version of face R, so that the monkey was never shown the unambiguous version of face R. By trial block 5, the response to the ambiguous version of face H was 72 spikes s^{-1}, whereas the firing rate to the ambiguous version of face R remained at 31 spikes s^{-1}. A one-way analysis of variance and *post hoc* Newman–Keuls tests showed that by trial block 5 there was a significant increase in the response to the ambiguous version of face H but not of face R compared with the response to the same images in trial block 1. A two-way analysis of variance where one factor was face H *vs* face R, and the other was firing in stimulus block 1 *vs* firing in stimulus block 5 showed that there was a significant interaction ($F[59,1] = 12.2$, $p < 0.001$) (neurone am240). The experiment thus indicates that seeing face H in its unambiguous form (in trial blocks 2 and 4) increased the response of the neurone to the ambiguous version of the same face (seen later in trial blocks 3 and 5), and that this effect was specific to the face that had been shown in the unambiguous form, in that the response to face R did not show a comparable increase.

$\ll 0.001$, as tested by Fisher's generalized significance test, which provided a value of 160.5 with 42 degrees of freedom. The fact that it was possible to demonstrate learning in one-third of the cells represents an underestimate, in that a proportion of the cells was not especially tuned to respond to the unambiguous stimulus.

In three of the experiments the control stimuli were ambiguous faces which were never shown in their unambiguous form. The system increased its response selectively to only the ambiguous face which had been shown previously in the unambiguous form. In four of the experiments the control stimuli were ambiguous non-face stimuli of which the non-ambiguous version was shown in trial block 2, and the system increased its response selectively only to the ambiguous face which had been shown in unambiguous form in trial block 2. In three of the experiments, one block of 10 trials with the unambiguous face (i.e. 5 s total experience with the stimulus) was sufficient to increase the response to the ambiguous version of the same face, and in four experiments two blocks of 10 trials (i.e. 10 s of experience) were needed.

Discussion

The results demonstrate that the responses of single neurones in the temporal cortical visual areas can become rapidly modified, showing effects of learning with as little as 5–10 s of experience of a complex visual stimulus in such a way that generalization occurs to much reduced representations of the same stimulus. The neuronal responses to the ambiguous image alter as a result of the experience even though the physical input to the neurone remains the same.

We also performed an experiment which showed that with similar stimuli to those used in the neurophysiological experiments presented under similar conditions humans also show rapid perceptual learning. We demonstrated this in a psychophysical experiment with 18 human subjects. The subjects were informed that something discernible or recognizable may appear in a visual display. Two fragmented (i.e. binarized or ambiguous) face stimuli (a+b) were first presented alternately 10 times for 0.5 s each. This was followed by a second block of trials in which face a was replaced by an easily recognizable version (a') (with 16 levels of grey scale), and face b was replaced by a grey scale version of a new face. Finally, in the last block the original fragmented faces (a+b) were once again repeated twice in alternation. The subjects never identified b as a face in the 36 trials, but accurately identified a as a face on 27 of the 36 trials. This difference is statistically highly significant (two-sample test for the difference between two proportions, $z = 6.57$, $p << 0.001$). Effects of similar learning can be demonstrated in humans 1 week after seeing stimuli analogous to the type used here[1] (A.Treisman, personal communication).

The experiments described here provide evidence that rapid learning about stimuli in the world is evident in the responses of neurones in the temporal cortical visual areas. This rapid learning may underlie the perceptual learning about new objects that enables us to recognize new objects or faces seen for only one or a few seconds previously. The findings are also consistent with current hypotheses that rapid learning should be a property of neurones in these temporal cortical areas, as part of a system for learning about the rapidly changing views of objects which enable us to recognize those objects later.[10,11] The rapid learning demonstrated here may be important in learning many other types of invariance, including size and position invariance, which are major computational problems that are solved by our visual system.[11-13]

Conclusion

A population of neurones in the primate temporal cortical visual areas shows an enhancement of response to an ambiguous stimulus after brief (e.g. 5 s) exposure to the corresponding unambiguous stimulus. This neuronal response modification may underlie the rapid perceptual learning to similar stimuli found in humans.

References

1. Ramachandran, VS. 2-D or not 2-D – that is the question. In: Gregory RL and Harris J, eds. *The Artful Eye*, Oxford: Oxford University Press, 1994.
2. Gross CG, Rocha-Miranda CE and Bender DB. *J Neurophysiol* **35**, 96–111 (1972).
3. Perrett DI, Rolls ET and Caan W. *Exp Brain Res* **47**, 329–342 (1982).
4. Desimone R. *J Cogn Neurosci* **3**, 1–8 (1991).
5. Perrett DI, Hietanen JK and Oram MW *et al. Phil Trans R Soc Lond B* **335**, 23–30 (1992).
6. Tovee MJ, Rolls ET, Treves A *et al. J Neurophysiol* **70**, 640–654 (1993).
7. Rolls ET and Tovee MJ. *Proc R Soc Lond B* **257**, 9–15 (1994).
8. Tovee MJ, Rolls ET and Azzopardi P. *J Neurophysiol* **72**, 1049–1060 (1994).
9. Rolls ET. *Phil Trans Roy Soc Lond B* **335**, 11–21 (1992).
10. Logothetis NK, Pauls J, Butthoff HH *et al. Current Biol* **4**, 401–414 (1994).
11. Rolls ET. *Behav Proc* **33**, 113–138 (1994).
12. Rolls ET. *Behav Brain Res* **66**, 177–185 (1995).
13. Wallis G and Rolls ET. *Prog Neurobiol*, in press (1996).

ACKNOWLEDGEMENT: This research was supported by the Medical Research Council, PG85135759. V.S.R. was a McDonnell-Pew Visiting Fellow at Oxford University. We thank Jonathan P. Stone for running the human psychophysics experiment described.

Received 4 June 1996; accepted 19 June 1996