The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in…

2 authors:

Edmund T Rolls
Oxford Centre for Computational Neuroscien…
702 PUBLICATIONS   **57,978** CITATIONS

SEE PROFILE

Martin J Tovée
University of Lincoln
164 PUBLICATIONS   **7,043** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project How can we help parents recognise unhealthy weight in their children? View project

Project Nicaragua Television and Beauty Ideals Project View project

ORIGINAL PAPER

Edmund T. Rolls · Martin J. Tovee

The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field

Abstract Neurons in the temporal visual cortical areas of primates have large receptive fields, which can show considerable selectivity for what the stimulus is irrespective of exactly where it is in the visual field. This is called translation invariance. However, such results have been found when there is only one stimulus in the visual field. The question arises of how the visual system operates in a cluttered environment. To investigate this we measured the responses of neurons with face-selective responses in the cortex in the anterior part of the superior temporal sulcus of rhesus macaques performing a visual fixation task. We found that the response of neurons to an effective face centred 8.5° from the fovea was decreased to 71% if an ineffective face stimulus for that cell was present at the fovea. In a similar way, introduction of a parafoveal ineffective face stimulus decreased the responses of these neurons to an effective face stimulus at the fovea to 75%. In addition to these interactions, it was found that an effective stimulus object at the fovea produced a larger response than when it was parafoveal, and that this weighting towards an object at the fovea was also seen when more than one object was present in the visual field. The implication of this weighting of the responses of neurons towards objects at the fovea, even in an environment with more than one object present, is that the output of the visual system provides information to subsequent systems particularly about objects at the fovea, so that learning about these objects (and less about other objects elsewhere in the visual field) is facilitated.

Key words Visual recognition · Attention Face identification · Translation in variance Monkey

E.T. Rolls (✉) · M.J. Tovee[1]
University of Oxford, Department of Experimental Psychology,
South Parks Road, Oxford, OX1 3UD, UK;
Fax no: +44-865-310447

Present address:
[1]Department of Psychology, University of Newcastle,
Ridley Building, Newcastle upon Tyne, NE1 7RU, UK

Introduction

There are two main processing streams: the dorsal stream, which is concerned with the relative spatial position of an object, and the ventral stream, which is concerned with its identification (Mishkin et al. 1983). Damage to the dorsal stream, through which the striate cortex projects to the parietal lobes, produces impairments on visuospatial tasks (Pohl 1973; Mishkin et al. 1982). Damage to the ventral system, through which the striate cortex projects to the temporal lobes, does not impair performance of visuospatial tasks but does impair performance of object discrimination tasks (Cowey and Gross 1970; Gaffan et al. 1986; Weiskrantz and Saunders 1984). As visual information passes through the ventral system a representation of objects is built up that is invariant with respect not only to the cues which define it, such as colour, motion or luminance (Sary et al. 1993), but also to the size of the object (Rolls and Baylis 1986). In contrast, cells in the primary visual cortex (V1), at the start of the processing stream, are much more finely tuned to size/spatial frequency, with a typical spatial frequency tuning of 1.5 octaves. Further, one population of cells in the cortex forming the ventral lip of the superior temporal sulcus has responses which are relatively view invariant (Hasselmo et al. 1989; Perrett et al. 1991). It is hypothesised that along the ventral processing stream representations of objects are built which are invariant with respect to, for example, size, rotation and view, for such invariant representations form appropriate inputs to associative neuronal networks in the hippocampus and amygdala which appear to be involved in an intermediate term "on the fly" buffer store and in stimulus-reinforcement association memory (see, e.g. Rolls 1990, 1992a, b, 1994a, b). Consistent with this, lesions of the inferior temporal visual cortex impair the ability of monkeys to respond to objects irrespective of changes in, for example, size, lighting and viewing angle (Weiskrantz and Saunders 1984).

It is likely that invariant representations for retinal position are also built along the same processing stream

(Gross and Mishkin 1977; Rolls 1992a, 1994a, b; Wallis et al. 1993). Consistent with this, the receptive fields of single cells in VI are usually small (e.g. 0.5–1°). In contrast the receptive fields of neurons in the inferior temporal visual cortex are very large, almost always including the fovea, and often extend to include large areas of the contralateral visual field (Gross et al. 1972; Desimone and Gross 1979). However, these studies on inferior temporal neurons were performed under anaesthesia and pharmacological immobilization, and it was not clear that the neuronal responses found will reflect those which occur when the cortex is operating normally, while the animal is awake and behaving. Recent studies on unanaesthetized macaques have shown that the responses of face-selective neurons in the lower lip of the superior temporal sulcus and the inferior temporal cortex were invariant with respect to the fixation position on the face, and typically also showed largely translation invariant but still stimulus-selective responses when the fixation position was 2–6° away from the edge of a face (Tovee and Rolls 1993; Tovee et al. 1994).

In the preceding investigations of translation invariance in the responses of temporal cortical neurons, the neuronal responses were measured when only one stimulus was placed at different positions relative to the fovea in the visual field. It is of considerable importance for understandig how the system operates normally in the visual world to ask how the system would respond when more than one object is present in the visual field, as would normally be the case. Would it be the case that if a non-effective stimulus was being fixated, but an effective visual stimulus for that cell was present anywhere in the visual field, the cell would respond? This might be expected by simple extrapolation from the findings already described on the translation invariance of single neurons in the temporal cortical areas. On the other hand, if the system did operate in this way, then it would mean that there would be massive activation of great numbers of inferior temporal neurons in normal, visually rich environments. In the case of face-selective cells, the result would be that a cell would respond whenever an effective face for a cell was present anywhere in the visual field. Alternatively, it might be the case that under conditions when many stimuli were present, a cell would show its full response to an effective stimulus only when that stimulus was fixated. Similarly, if a non-effective stimulus were present at the fovea, but an effective stimulus were present parafoveally, then the cell might not respond maximally to its effective (parafoveal) stimulus. Operation in this latter way would have a number of advantages, including simplifying the interface to the memory systems which receive outputs from the visual system, for then there would be much less clutter in the output, which would be weighted towards the visual stimulus being fixated. Such a better-defined output would also be useful for an interface to action and motor systems, for then the stimulus to which action should be directed would not be ambiguous.

Some interactions between simple visual stimuli presented in different parts of the visual field have been described previously for neurons in the inferior temporal visual cortex (area TE) of macaques. Sato (1989) showed that neurons that responded to stimuli such as bars or coloured spots typically showed a similar or smaller response when two stimuli were present than when the better of the stimuli was presented alone. (The neurons he studied frequently did not show very good translation invariance, and the effects were stronger for stimuli presented contralaterally than ipsilaterally). Miller et al. (1993), using simple visual stimuli based on Fourier descriptors, also found that the responses of most of the inferior temporal cortex (TE) neurons they studied were weaker to pairs of such stimuli than to the better single stimulus of the pair presented alone. Sato (1988), in line with earlier findings of Richmond et al. (1983), found that attention directed to a stimulus in order to detect its dimming reduced the responses of inferior temporal cortex (area TE) neurons to simple visual stimuli such as a bar of light. In the study described here, we pursued the investigation beyond interactions between simple stimulus features such as bars or coloured spots and focused on interactions that occur when more than one complex object (such as a face, itself consisting of many features), it present in the visual field. Indeed, in the study described here, we specifically investigated the responses of neurons with face-selective responses, in order to investigate how the visual system codes for complex stimulus objects of this type when more than one object (face) is present in the visual field. Moreover, in this study we investigated the responses of neurons in the cortex in the anterior part of the superior temporal sulcus, in which there have been no previous investigations of interactions between more than one object or stimulus present in the visual field. In the study described here we also made it a point not to require focal attention of the monkey to one or other of the stimuli, in order to approximate viewing conditions of the real world which occur frequently. We also took care to ensure that a fixation spot did not interfere with the neuronal responses to the test stimuli (see Richmond et al. 1983), by using a blink version of a visual fixation task. The properties of these neurons with face-selective responses have been described in a number of studies (see Rolls 1992a), and they are appropriate for the investigation described here also because a number of their properties have been analysed in detail (see Rolls 1992a; Desimone 1991; Rolls 1991).

The plan of the experiments described here was to measure the responses of inferior temporal neurons when a non-effective face stimulus for a cell was present at the fovea, to determine whether this affected the otherwise large response to an effective face presented parafoveally, and to determine whether the response to an effective face stimulus presented at the fovea was altered by non-effective face stimuli presented parafoveally. The awake, behaving monkeys performed a visual fixation task, and eye position was measured with the search coil technique, to ensure that the stimuli were presented in known positions in the visual field.

This investigation is part of a series of experiments performed in order to understand the normal functioning of the cerebral cortex and how disorders of its function lead to perceptual and cognitive dysfunctions (Rolls 1992a, 1994a, b; Tovee and Rolls 1992; Tovee et al. 1993, 1994; Rolls and Tovee 1994; Rolls et al. 1994a, b;).

Materials and methods

Recording techniques

The activity of single neurons was recorded with glass-insulated tungsten microelectrodes (after Merrill and Ainsworth 1972, but without the platinum plating) in three alert macaque monkeys (*Macaca mulatta*, weight 3 kg) seated in a primate chair using techniques that have been described previously (Rolls et al. 1976, 1990; Tovee et al. 1993). The action potentials of single cells were amplified using techniques described previously (Rolls et al. 1979), were converted into digital pulses using the trigger circuit of an oscilloscope and were analysed on line using a MicroVaxII computer. The computer collected peristimulus rastergrams of neuronal activity for each trial and displayed, printed and stored each trial, as well as computing the peristimulus time histogram by summing trials of a given type. Eye position was measured to an accuracy of 0.5° with the search coil technique (Judge et al. 1980), and steady fixation of a position on the monitor screen was ensured by use of a blink version of a visual fixation task in which the fixation spot was blinked off 100 ms before the target stimulus appeared. The timing of the task is described below. The stimuli were static visual stimuli subtending 8.5° in the visual field presented in the top left and/or top right quadrant of the video monitor, placed at a distance of 1.0 m from the eyes. The fixation spot position was at the middle of either the top left or the top right quadrant of the screen. The angle between the centre of the two quadrants was 8.5° (see Fig. 2). The monitor was viewed binocularly, with the whole screen visible to both eyes.

Task timing

Task timing is shown in Fig. 1. Each trial started at −500 ms (with respect to the onset of the test image) with a 500-ms warning tone

to allow fixation of the fixation point, which appeared at the same time. At −100 ms the fixation spot was blinked off so that there was no stimulus on the screen in the 100-ms period immediately preceding the test image. The screen in this period, and at all other times including the inter-stimulus interval, was set at the mean luminance of the test images and the mask. At 0 ms, the tone was switched off and the test image was switched on for 500 ms. At the termination of the test stimulus the fixation spot reappeared, and then after a random interval in the range 150–3350 ms it dimmed, to indicate that licking responses to a tube in front of the mouth would result in the delivery of saline. The dimming period was 500 ms and after this the fixation spot was switched off, and reward availability terminated 500 ms later. The monkey was required to fixate the fixation spot in that, if he licked at any time other than when the spot was dimmed, saline instead of fruit juice was delivered from the tube, and in that, if the eyes moved by more than 0.5° from time zero until the start of the dimming period, then the trial was aborted. (When a trial aborted, a high-frequency tone sounded for 0.5 s, no reinforcement was available for that trial, and the inter-trial interval was lengthened from 8 to 11 s). It should be noted that this is not a task in which attention either at or away from the test visual stimuli was explicitly rewarded. The reward the monkey received was not contingent on the monkey making any response to the test stimuli. What the fixation task does is to allow the fixation position of the monkey to be controlled. This then enabled stimuli to be shown in different positions with respect to the fovea. In this way we were able to examine the effects of having more than one visual stimulus present in the visual field at one time.

Stimuli

Both macaque and human faces were used as test stimuli. Examples have been shown by Baglis et al. (1985) and Leonard et al. (1985). The stimuli were stored in digital form on a computer disk, and displayed on a monochrome video monitor using a video framestore (Advanced Electronic Design 512). The resolution of these images was 256 wide by 256 high with 256 gray levels. At all times the mean luminance of whatever was displayed on the monitor (even when it was a blank screen) was set to gray level 128, so that there were no luminance changes during each trial. The monitor provided maximum and minimum luminances of 6.0

Fig. 1 The timing used in the visual fixation blink task

412
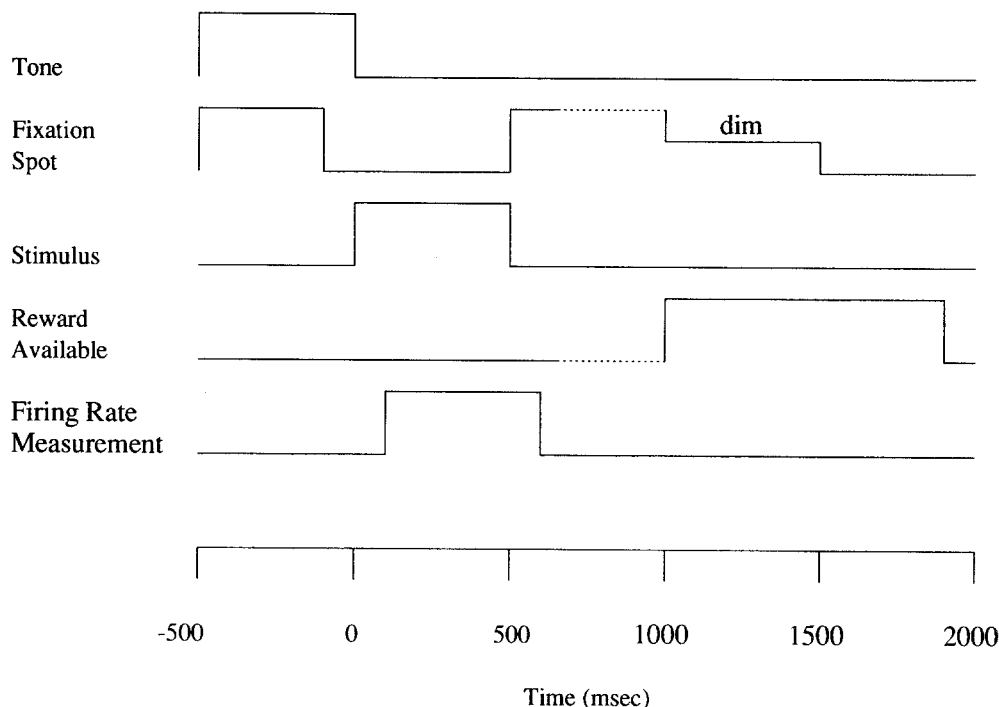
and 0.13 footlamberts (ftl), respectively, and was adjusted internally and by use of a lookup table for linearity to within 3% using a photometer. The computer randomized the order of presentation of the test stimuli, switched the stimuli on and off for each trial and synchronized its data collection so that the stimulus was turned on at the start of the 21st bin of the peristimulus time histogram.

During initial investigation of the responsiveness of a temporal cortical cell, when digitized visual stimuli were being presented on the video monitor, one set of 4–12 visual stimuli were used at a time. Each set of stimuli was designed to provide neuronal response data relevant to one or several hypotheses. For example, one set included five different faces, to test whether the neuron responded differently to different faces, and some non-face stimuli such as a sine wave grating, a boundary curvature descriptor, and a complex visual image (see Baylis et al. 1985, their Fig. 1), to provide an indication of whether the neuron responded differently to face and to non-face stimuli. The computer randomized the sequence in which the members of the set were presented, and after it had presented the sequence once, it restarted the set with another random sequence. The computer was allowed to repeat the set 4–10 times in order to provide sufficient data for an analysis of variance, so as to determine whether the neuron responded differently to the different stimuli within the set.

Procedure

As tracks were made into the cortex in the superior temporal sulcus, the responses of each neuron were measured to a standard digitized set of stimuli of different faces and of non-face stimuli (Baylis et al. 1985). If a neuron responded to one or more of the faces, but to none of the non-face stimuli in the set, then a wide range of digitized and real three-dimensional (3D) non-face stimuli were shown, to determine whether the response of the neuron was selective for faces. The criterion was that the response to the optimal face stimulus should be more than twice as large as to the optimal non-face stimulus. (In fact, the majority of the neurons in the cortex in the superior temporal sulcus classified as showing responses selective for faces responded much more specifically than this. For half these neurons, their response to the most effective face was more than 5 times as large as to the most effective non-face stimulus, and for 25% of these neurons the ratio was greater

than 10:1. These ratios show that, while responding preferentially to faces, these neurons do not have absolute specificity for faces. Further information on and discussion of the extent to which these neurons have selective responses is given by Baylis et al. (1985) and by Rolls (1984, 1992a, c). The non-face stimuli from which the optimal was chosen included sine wave gratings, boundary curvature descriptors, complex 2D stimuli and complex 3D junk objects, as described above.)

If the neuron satisfied the criterion for face selectivity, then the following experiment was performed to assess how the presence of more than one visual stimulus in the receptive field of the neuron affected the neuronal responses. Two or more face stimuli were selected, usually including one effective and one relatively ineffective in driving the cell. One of the two fixation positions shown in Fig. 2 was selected at random and then held constant for the remainder of the experiment on that neuron. Then three test conditions were run. These were presentation of stimulus 1 at the fixation point, presentation of stimulus 2 in the non-fixated quadrant, or presentation of both stimuli simultaneously in the same quadrants in which they were presented for the first two conditions. The order of these conditions in a test sequence was randomised. Once the set of three stimuli had been presented once in random order, the set was repeated with a new random order and testing was continued until 6–12 sequences had been completed. This provided six to ten repetitions for each stimulus condition.

This experimental design was used in the following ways. First, the non-effective face stimulus (for the particular cell being recorded throughout) could be allocated to the quadrant being fixated and the effective face stimulus to the quadrant not being fixated. This allowed testing of the effects of a non-effective stimulus at the fixation point on what would without it be a large neuronal response to the parafoveally presented effective visual stimulus. Second, the effective face stimulus could be allocated to the quadrant being fixated and the non-effective face stimulus to the quadrant not being fixated. This allowed testing of the effects of a non-effective stimulus parafoveally on what would without it be a large neuronal response to the effective visual stimulus presented at the fovea. Every cell described here was tested in both the first and the second conditions.

Third, the effective face stimulus could be allocated to the quadrant being fixated, and another different effective face stimu-

Fig. 2A, B The arrangement of the stimuli and fixation spots on the screen of the video monitor. The stimuli were normally placed in the upper left and/or upper right quadrant of the screen, and the fixation point was at the centre of one of these quadrants. The angle subtended by one of the images was 8.5°. The effective stimulus for a cell is depicted with *dark shading* and the ineffective stimulus with *light shading*. **A** The three test conditions for investigation of whether the response to a parafoveal effective face stimulus was affected by the presence of an ineffective visual stimulus at the fixation point. **B** The three test conditions for investigation of whether the response to a foveal effective face stimulus was affected by the presence of an ineffective visual stimulus parafoveally
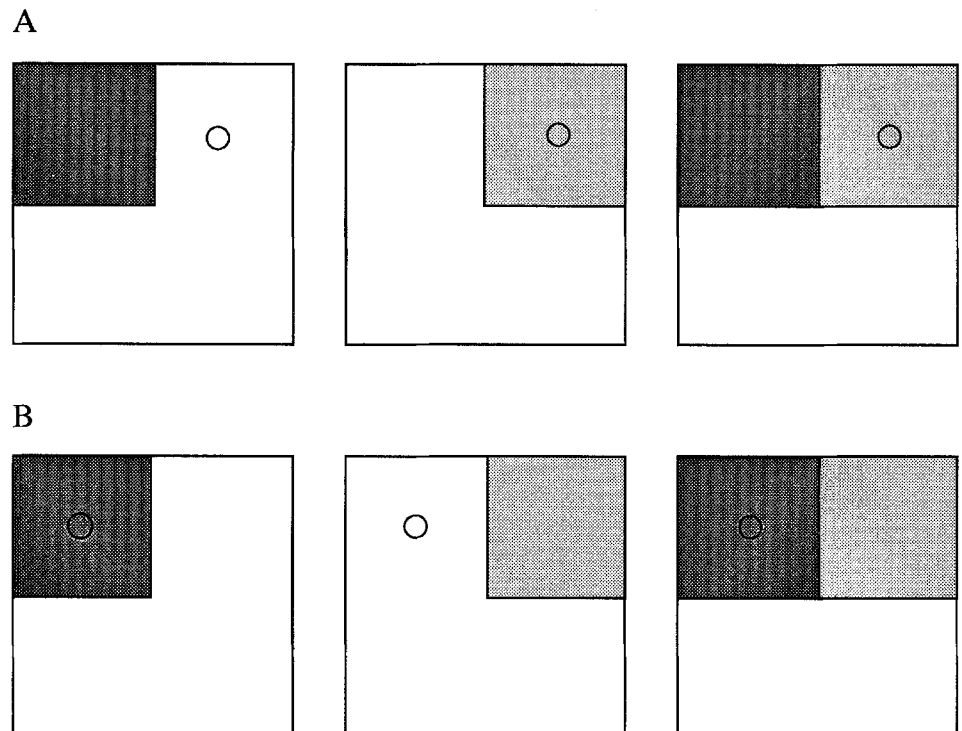
lus to the quadrant not being fixated. This allowed testing of whether there would be linear summation, or any type of interaction, between two effective visual stimuli present simultaneously in the visual field. Fourth, we also investigated whether there were interactions, or linear summation, between two different non-effective face stimuli for a particular cell present simultaneously.
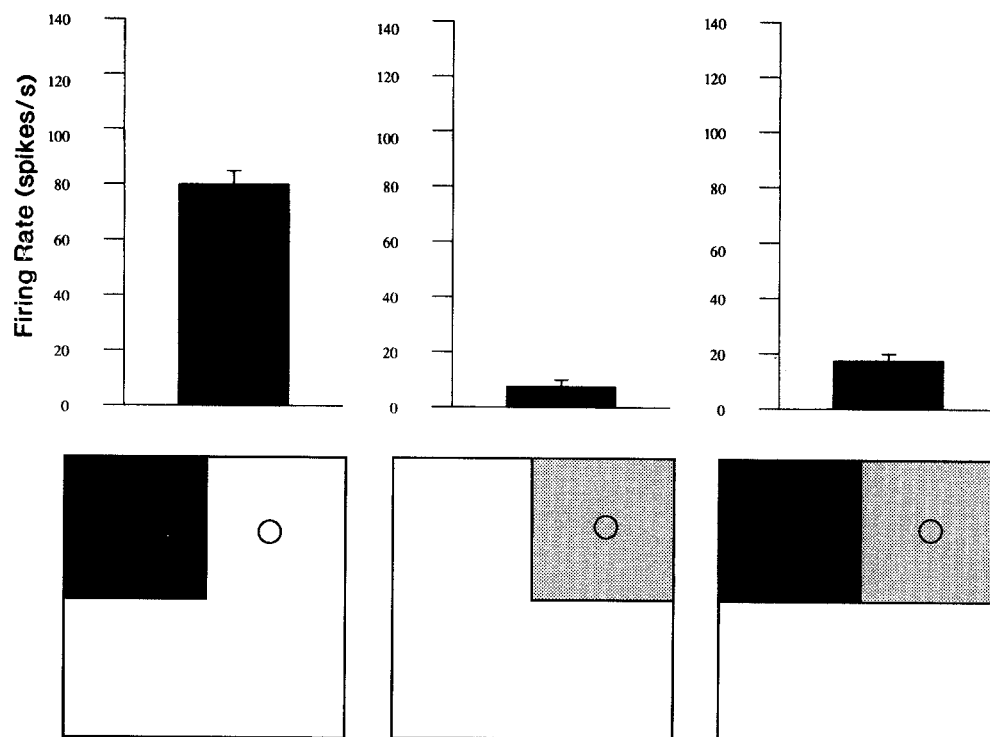
The trials in this testing protocol followed each other with a fixed inter-trial interval, and all the testing and neuronal spike collection were fully automatic. All the neurophysiological data were saved to disc for subsequent analysis.

Data analysis

The default period for which the firing rate was calculated was a 500-ms period starting 100 ms after the onset of the target stimulus. The period was chosen to start at 100 ms because all the neurons had started to respond strongly by this time. (Typical response latencies for the neurons were 70–90 ms). The mean neuronal response over the 6–12 trials for each stimulus condition together with its standard error was calculated by the computer for each of the stimulus conditions for graphical presentation. The neuronal response was calculated at the firing rate in the specified period minus the mean spontaneous firing rate of the neuron in the interstimulus interval. The mean across all the neurons described here of the spontaneous firing rate was 9.3 spikes/s (SD 5.6). (Throughout this paper, firing rate indicates the absolute firing rate of the neuron, and the response indicates the firing rate minus the spontaneous rate of the neuron.) In addition, a one-way analysis of variance was performed over the same set of data, with the three conditions described above being compared; t-tests using the appropriate error term from the analyses of variance and Tukey tests were then used if particular pairs of neuronal responses were to be compared.

Peristimulus rastergrams were produced to show the temporal aspects of neuronal firing, and peristimulus time histograms were made from the spike arrival times, which were saved with an accuracy of 1 ms. A peristimulus spike density function was computed from the 1-ms time histogram formed from 6–12 trials of a given stimulus condition smoothed with a Gaussian filter (with a σ of 10 ms and a cutoff of 3 SDs). These functions confirmed that 100–600 ms poststimulus was a relevant response period for measurement of the neuronal firing.

Recording sites

X-radiographs were used to locate the position of the microelectrode on each recording track relative to permanently implanted reference electrodes and bony landmarks. The position of cells was then reconstructed from the X-ray co-ordinates taken together with serial 50-μm histological sections which showed the reference electrodes and micro-lesions made at the end of some of the microelectrode tracks (Feigenbaum and Rolls 1991).

Results

The responses of a large sample of cells in the temporal cortical visual areas were recorded on both sides of the brain. It was possible to perform 96 experiments on interactions of stimuli in the receptive field on 33 neurons in three monkeys. An experiment consisted of one of the four stimulus configuration paradigms described in the Materials and methods section. Whenever paradigm 1 (a parafoveal effective stimulus and a foveal noneffective stimulus) was run, paradigm 2 (a foveal effective stimulus and a parafoveal noneffective stimulus) was also run on the same cell (with a different order for different cells), so that the magnitudes of the interactions could be directly compared for each cell.

An example of the results of an experiment in which an effective face stimulus was shown parafoveally is shown in Fig. 3. The fixation position was at the point indicated by a dot in the centre of the upper right quadrant on all trials. When only the effective face stimulus was shown parafoveally (in the upper left quadrant), the neuron had a large response to it (Fig. 3, left). The magnitude of the response in this condition was almost the same as when the effective stimulus was at the fovea.
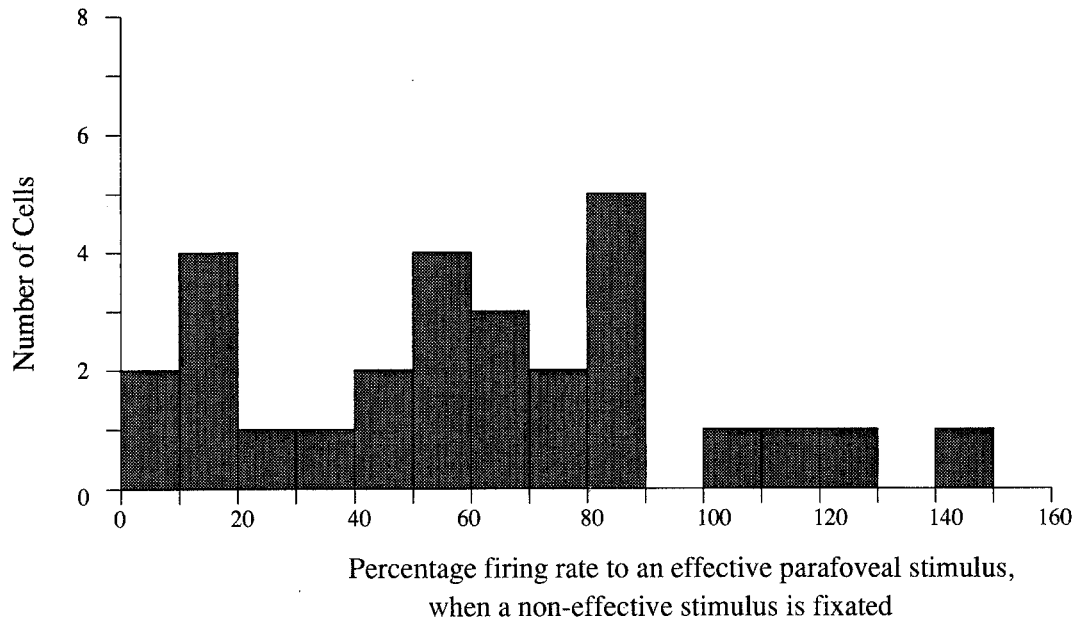

Fig. 3 The neuronal firing rate of one cell when an effective stimulus was present parafoveally and an ineffective stimulus for that cell was present at the fovea. Fixation was always at the point shown by a *dot* in the centre of the upper right quadrant. *Left*, the neuron had a large firing rate when the effective stimulus was shown parafoveally. *Middle*, the neuron did not respond when the non-effective stimulus was present at the fovea. *Right*, the neuron had only a small firing rate to the parafoveal effective stimulus if a non-effective stimulus was present at the fovea. The means and standard errors of the firing rate are shown. The mean spontaneous rate of the cell was 10 spikes/s

414

Percentage firing rate to an effective parafoveal stimulus,
when a non-effective stimulus is fixated

Table 1 Mean responses of neurons in the macaque temporal visual cortex to effective and non-effective stimuli

	Effective alone	Non-effective alone	Both
Fixation on non-effective (spikes/s)	34.6	12.6	22.4
Fixation on effective (spikes/s)	49.0	11.6	34.4

Fig. 4 The neuronal responses of the population of cells when an effective stimulus was present parafoveally and an ineffective stimulus for that cell was present at the fovea. The histogram shows the percentage of the cells's response to an effective parafoveal face that was produced when in addition there was an ineffective face at the fovea

This is the basic finding of translation invariance, which was found for all the cells described in this paper, and has been more fully documented by Tovee et al. (1994). The fact that this cell did not respond to the non-effective visual stimulus is shown in Fig. 3 (middle). The stimulus was shown in the upper right quadrant, where the monkey was fixating. The interesting test condition is on the right in Fig. 3. The effective face stimulus was shown parafoveally, and the cell responded to it very much less than previously as shown on the left in Fig. 3, associated with the presence of a non-effective stimulus where the monkey was fixating. (The responses of the cell to the three stimulus conditions was significantly different (one-way ANOVA, F=207.34 df=2, 117, P<0.00001), and the response to the effective face shown parafoveally is significantly different from the simultaneous presentation of both faces (Tukey test, P<0.01). Thus for this cell, translation invariance did not operate irrespective of what other stimulus was in the receptive field. Instead, the presence of a non-effective stimulus at the fovea decreased the response of the neuron to an effective stimulus that was presented parafoveally.

Some reduction was found for the majority of the 28 neurons analysed with this arrangement of stimuli, as shown in Fig. 4. In Fig. 4 the percentage is shown for each cell of the response to a parafoveal effective stimulus alone which occurred when there was an ineffective stimulus at the fovea. For 82.5% of the cells shown in

Fig. 4, the reduction in the response was statistically significant (65% of the cells showed a decrease significant at the P<0.01 level and 17.5% of the cells showed a decrease that was significant at the P<0.05 level). The mean across the 28 cells of the percentage of the response (firing rate spontaneous) with both stimuli compared with that with only the effective stimulus parafoveally was 71.4%. The mean response rate to the effective stimulus presented parafoveally was 34.6 spikes/s, to the non-effective stimulus at the fovea was 12.6 spikes/s, and to both stimuli present simultaneously (with fixation on the non-effective stimulus) was 22.4 spikes/s (see Table 1). (Although the mean response to both faces was intermediate between that to the effective and the non-effective face, as shown in Table 1, individual cells showed effects of differing magnitude, as shown in Fig. 4, and the magnitude of these effects was not simply related to the size of the response to the effective and non-effective faces).

The results of the second type of experiment are illustrated in Fig. 5. Figure 5 (left) shows that the cell had a large response when an effective face stimulus for the cell was presented at the fovea. (In this experiment, the fixation position was in the centre of the upper left quadrant of the screen). There was little response when the non-effective face stimulus for that cell was presented parafoveally, as shown in Fig. 5 (middle). However, the firing rate of the neuron when the monkey was looking at the effective face stimulus was decreased from 110 to 37 spikes/s if the ineffective stimulus was present parafoveally in the visual field (Fig. 5, right). The cell's response to the three stimulus conditions was significantly differ-

Fig. 5 The firing rate of the same cell as that shown in Fig. 3 when an effective stimulus was present at the fovea and an ineffective stimulus for that cell was present parafoveally. Fixation was always at the point shown by a *dot* in the centre of the upper left quadrant. *Left,* the neuron had a large firing rate when the effective stimulus was at the fixation point. *Middle,* the neuron did not respond when the non-effective stimulus was present parafoveally. *Right,* the firing rate of the neuron to the effective stimulus at the fixation point was reduced by the non-effective stimulus present parafoveally. The means and standard errors of the firing rates are shown. The mean spontaneous rate of the cell was 10 spikes/s
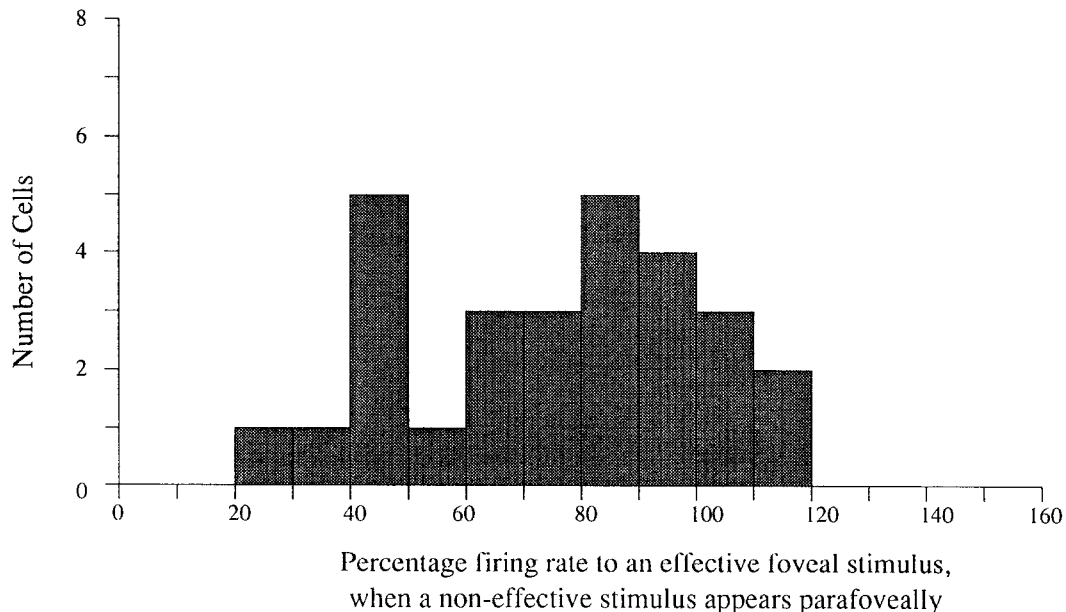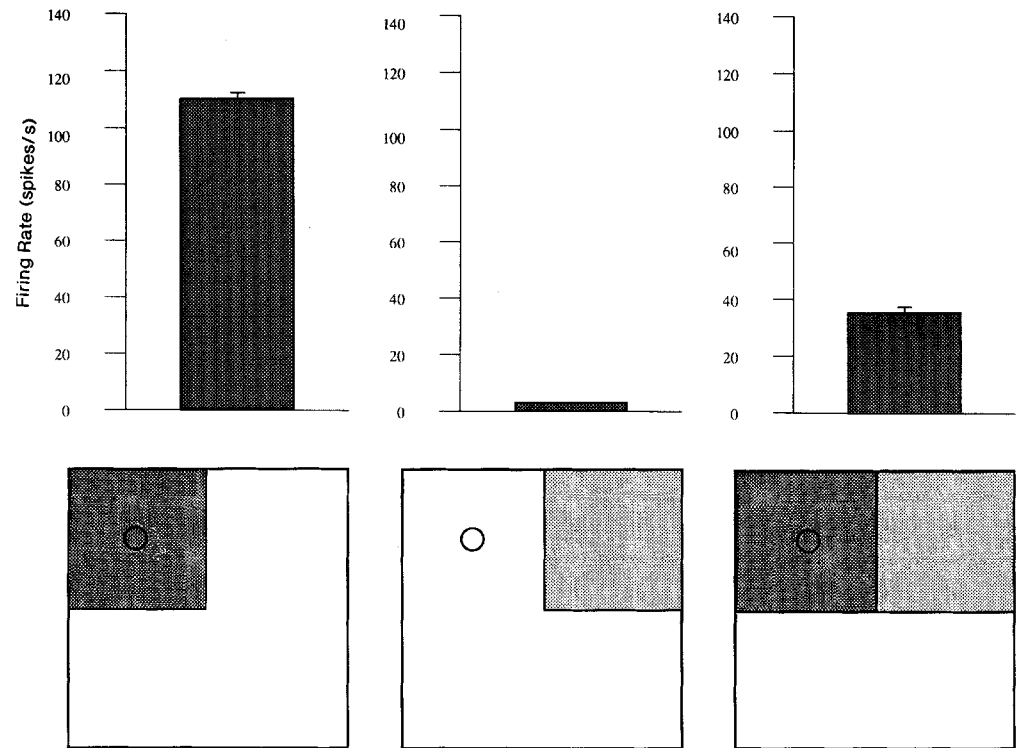

Percentage firing rate to an effective foveal stimulus, when a non-effective stimulus appears parafoveally

Fig. 6 The neuronal responses of the population of cells when an effective stimulus was present at the fovea and an ineffective stimulus for that cell was present parafoveally. Histogram in which the abscissa shows the percentage of the cell's response to an effective face that was produced when in addition there was an ineffective face for that cell parafoveally

ent (one-way ANOVA, $F=547.89$, $df=2,117$, $P<0.00001$), and the decrease in the cell's firing rate to the effective stimulus when a non-effective stimulus appeared parafoveally was statistically significant (Tukey test, $P<0.01$).

A similar result was found for the majority of the neurons tested, as shown in Fig. 6. For 60.0% of the cells

tested, the decrease in the neuronal response to an effective fixated face stimulus produced by the presence of a non-effective face stimulus was significant at $P<0.01$; for an additional 6.7% of the neurons, the decrease was significant at $P<0.05$. The mean percentage response rate was decreased to 75%. The mean response rate to the effective stimulus presented foveally was 49.0 spikes/s, to the non-effective stimulus parafoveally was 11.6 spikes/s and to both stimuli present simultaneously (with fixation on the effective stimulus) was 34.4 spikes/s (see Table 1).

The results for the third type of experiment are illustrated in Fig. 7. On the left (Fig. 7), it is shown that the neuron had a good response to one effective face stimu-

416

Fig. 7 The firing rate of the same cell as that shown in Figs. 3 and 5 when an effective stimulus was present at the fovea and another effective stimulus for that cell was present parafoveally. Fixation was always at the point shown by a *dot* in the centre of the upper left quadrant. *Left,* the neuron had a large firing rate when the first effective stimulus was parafoveally. *Middle,* the neuron had a similarly large firing rate when another effective stimulus was present foveally. *Right,* the firing rate of the neuron to the first effective stimulus parafoveally and to the second effective stimulus at the fixation point was the same as that to either stimulus separately. The means and standard errors of the firing rate are shown
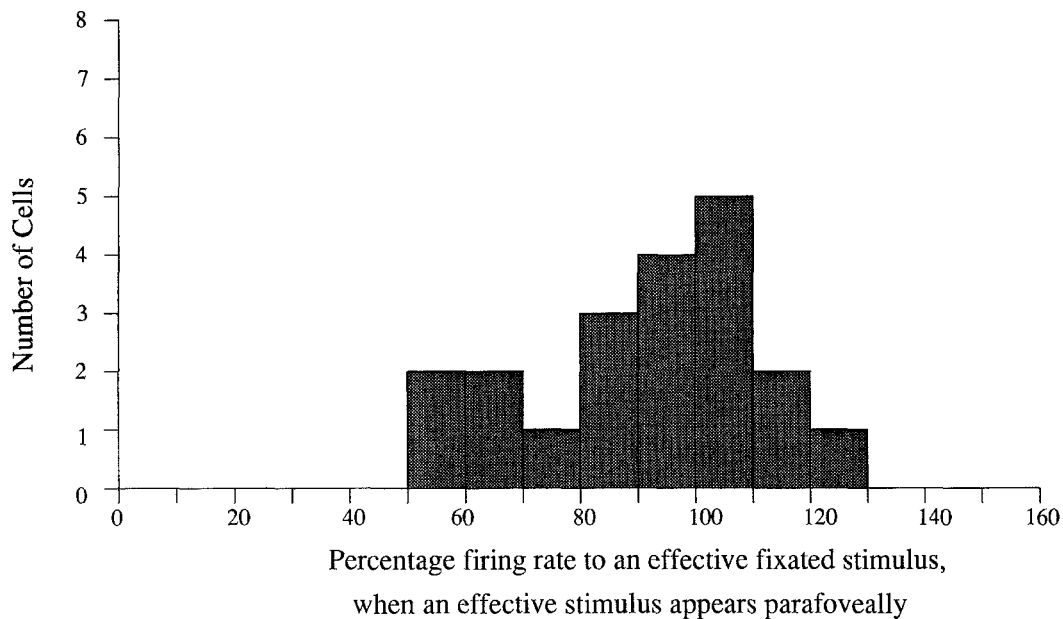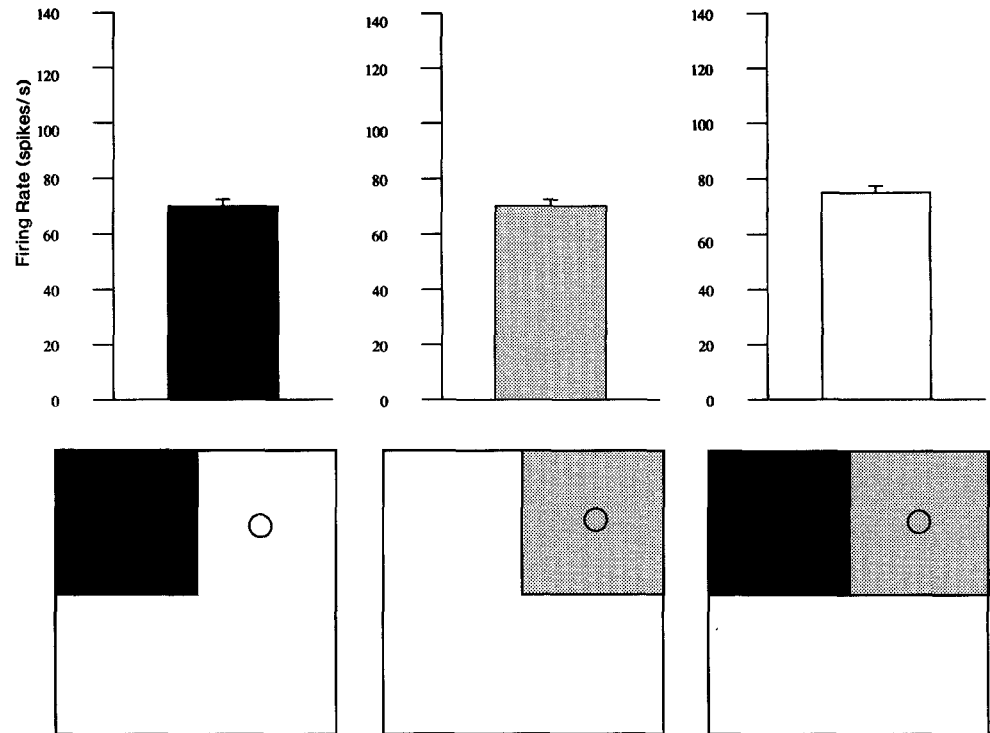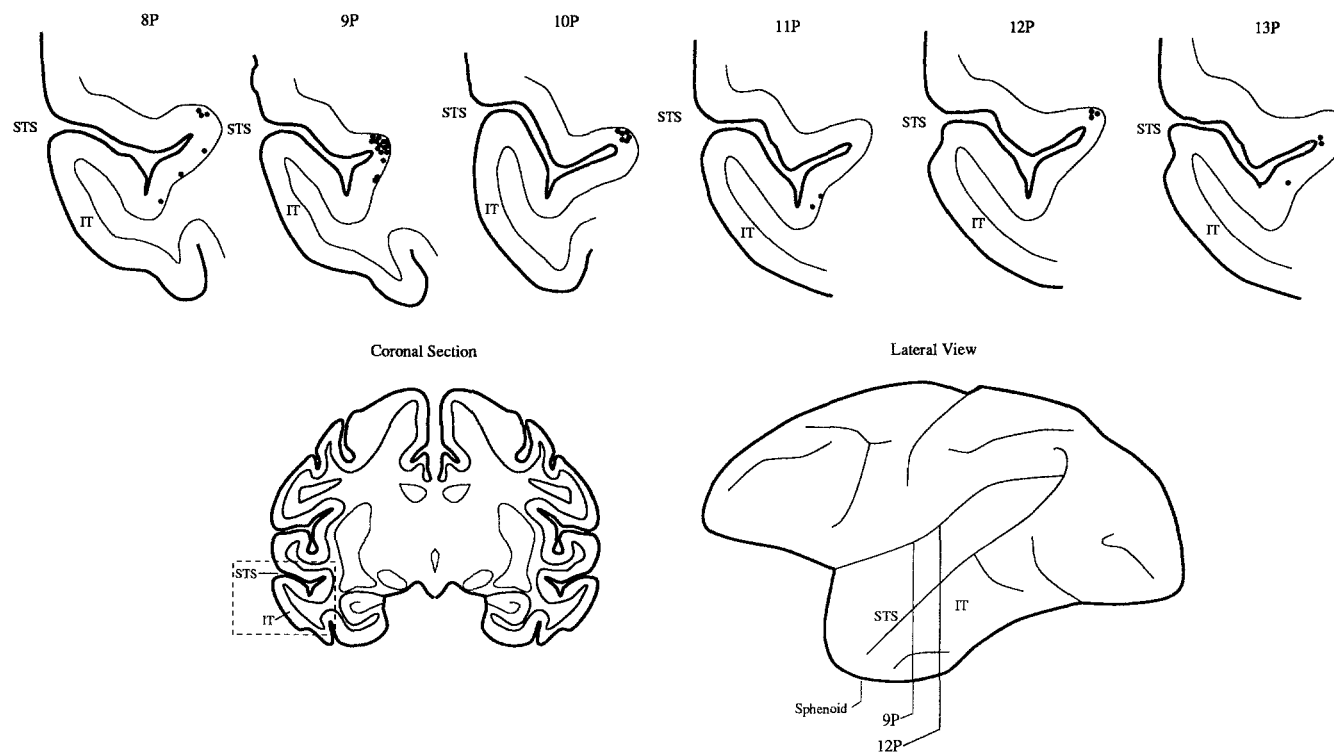

Percentage firing rate to an effective fixated stimulus, when an effective stimulus appears parafoveally

Fig. 8 The neuronal responses of the population of cells when an effective stimulus was present at the fovea and another effective stimulus for that cell was present parafoveally. Histogram in which the abscissa shows the percentage of the cell's response to an effective face at the fovea that was produced when in addition there was another effective face for that cell parafoveally

lus when it was presented alone parafoveally. Figure 7, middle, shows that the neuron had a comparably large response when another effective face stimulus was presented at the fovea. On the right, Fig. 7 shows that when both stimuli were presented simultaneously, the neuron

had a response which was not significantly different to that when either stimulus was presented alone (one-way ANOVA, $F=0.35$, $df=2$, 117, $P<0.711$). A similar result was obtained in the fourth experimental paradigm in which two non-effective face stimuli were used instead of the two effective stimuli.

Similar results were obtained for the majority of the neurons tested with two effective face stimuli. This is shown in Fig. 8, which makes it clear that most of the neurons' responses to an effective face presented at the fovea were little affected by presentation of another effective face parafoveally, even though that face presented

Fig. 9 The recording sites, shown on coronal sections, of the neurons included in this study. The positions of the coronal sections are shown on a lateral view of the macaque brain. The distances refer to millimetres posterior to the sphenoid reference plane (see text) (*STS* superior temporal sulcus, *IT* inferior temporal visual cortex)

parafoveally alone was capable of driving the cell well. Only 11% of the cells showed a statistically significant ($P<0.05$) change in firing rate to simultaneous presentation of the effective stimuli compared with presentation of the stimuli separately. The mean percentage response rate was decreased to 93%.

The recording sites of the neurons analysed in this study are shown in Fig. 9. All the cells analysed were in the cortex in the depths of the anterior part of the superior temporal sulcus.

Discussion

These neurophysiological results show that there are interactions between different object stimuli (faces) presented in different parts of the receptive fields of neurons in the visual areas of cortex in the anterior part of the superior temporal sulcus of the macaque. Translation invariance does not operate to produce responses to stimuli when more than one stimulus is present which reflect the magnitude of the response when only one stimulus is present. These findings are, we believe, fundamental to understanding how the visual system functions in a normal, cluttered, environment. The following issues are of especial interest.

First, translation invariance ceases to result in a large response to a parafoveal stimulus if there is another, relatively ineffective, face stimulus at the fovea (Figs. 3, 4). The mean response to the parafoveal effective stimulus (34.6 spikes/s) was reduced to 22.4 spikes/s when a non-effective foveal stimulus was present (see Table 1). The mean percentage reduction across cells was 71.4%.

Second, there was also a comparable reduction in the response to an effective face at the fovea (49 spikes/s) when a non-effective face was introduced parafoveally, to 34.4 spikes/s (see Table 1). The mean percentage reduction across cells was 75.4%. Thus the reduction in the response to an effective face produced by a non-effective face was approximately similar (in terms of percentage reduction), whether or not it was the effective face that was fixated.

Third, the response of the system when more than one object is introduced into the field of vision is clearly non-linear (non-additive). For example, in the first and second paradigms the neuronal response to both stimuli shown together did not reflect the addition of the responses to the two objects presented separately (in fact, there was a reduction in the response to the better stimulus). Further, in the third paradigm, when two effective stimuli were shown simultaneously, the neuronal response was not greater than to either stimulus presented alone (in fact, the neuronal response to both was on average just less than to either stimulus presented alone).

Taking the data as a whole shown in Table 1, the following points can be made about how this part of the visual system operates with more than one object present in the visual field. Although considerable translation invariance is shown by this set of neurons, the response to an effective stimulus was a little greater when the object was fixated (49 spikes/s) than when it was present para-

foveally (34.6 spikes/s). (Note that this result is not inconsistent with the findings of Tovee et al., 1994, for although there was apparently less decline with fixation away from the centre of an object, see their Fig. 7), the stimuli were in that experiment twice the size of those used in the experiments described here. With the parafoveal position of the stimuli used here, the nearest part of the parafoveal image to the fovea was 4.25°). Given the interaction effects shown in Table 1, it is interesting to note that even when two different objects (faces) are present in the visual field, then the output of cells that respond best to one of the objects is greater when that object is fixated than when it is situated outside the fovea. This indicates a way that the visual system may use to disambiguate its output. Instead of all cells tuned to different stimuli responding with an output which is fully translation invariant, so reflecting only whether that stimulus is present in the field of view, the visual system operates under cluttered as well as uncluttered conditions to give stronger weight to what is at the fovea. This means that the visual system provides an output which can be interpreted by subsequent systems (such as the amygdala and hippocampus) without the necessity for a full solution of the binding problem. The visual system appears to help subsequent structures by weighting the output to what is at the fovea. This makes it possible for subsequent structures to associate for example a fixated object with reward, but not to associate all other objects in the field of view with reward, because the response for objects at the fovea is considerably larger than for other objects present simultaneously elsewhere in the field of view. The cells in the output of the visual system which respond most in a cluttered world are those which are tuned to respond not only to the particular objects present in the scene, but especially strongly to those objects if they are at the fovea. This is what is shown in Table 1. We note that further clarification of the output of the neural system would be produced by eye movements, for the data shown in Table 1 indicate that in a cluttered environment neurons tuned to a particular face have their greatest response when that face is at the fovea.

The suggestion then based on these findings is that the visual system uses this weighting of objects at the fovea as an efficient way of passing an output from a sensory system to a memory or motor system. If instead the whole scene were analysed simultaneously, then a multiplicity of reinforcement signals would have to be decoded and bound (linked) correctly to all the objects in the scene. The motor (or some other) system would then have the task of parsing all object/reinforcement linked pairs and then producing appropriate actions. By making one object in the environment the main subject of neuronal firing, a large part of the otherwise major computational problem of dealing with multiple objects simultaneously (von der Malsburg 1990) is avoided (Rolls 1992, Ballard 1993) at some cost implied by where one is looking (or paying attention), a serial process, in processing time. We note that, given the expansion of the fovea in the cortical representation of vision earlier on in cortical

processing (Rolls and Cowey 1970; Cowey and Rolls 1975), no further special mechanism would be needed to weight the firing of temporal cortex neurons towards stimuli present at the fovea than parafoveally: the number of afferents from foveal regions reaching the temporal cortical visual areas would be greater than from parafoveal regions.

Another interesting aspect of the findings described here is that the interactions between different stimuli in the receptive field were an important determinant of cell firing even when the monkey was not given instructions to pay attention selectively to either of the stimuli. What the monkey was being required to do was to fixate a position on the screen. While he was fixating that position, and if there was no stimulus present at the fovea, the neuron responded to a parafoveal effective stimulus. If the monkey was performing the identical task, but there was a non-effective stimulus where the monkey was fixating, then the cell had a much smaller response to a parafoveal, otherwise effective stimulus (Figs. 3, 4). Thus the interaction effects were obtained during general viewing of a position in the world, in the absence of specific instructions to the monkey to require him to pay attention to a particular position in space. Of course, spatial attention might well be used to increase the selectivity of the processes described here and to help to deliver an unambiguous output to the brain systems following the temporal cortical visual areas. It has been shown, for example, that when monkeys perform a delayed match-to-sample task with reward given for only one of two stimuli present simultaneously, then the response of V4 neurons depends on which stimulus the monkey is rewarded for paying attention to (Moran and Desimone 1985). When Moran and Desimone recorded in inferior temporal cortex they found a similar, although weaker attention effect. If we had used an attention task we presumably would have found a similar result, and of course using spatial attention to enhance responses in certain parts of the visual field is a useful mechanism. However, in normal vision we do not usually focus on one particular object or pattern in our visual surroundings away from the fovea, but instead have a stimulus at the fovea, but also much visual clutter in the rest of the scene. We therefore feel that a task which does not involve such an attention component provides a more normal view of how the visual system usually operates. As previously mentioned, the task used here did not involve attention being controlled by reward towards one or other of the stimuli, in that the reward the monkey received was not contingent on the monkey making any response to the test stimuli.

The results described here are consistent in general with findings on the effects of interactions between simple stimuli found with inferior temporal cortex neurons (Richmond et al. 1983; Sato 1988, 1989; Miller et al. 1993), but extend those findings to the case where there are separate whole complex objects (in this case faces) in different parts of the visual field, revealing how responses to these different objects interact, and extend the find-

ings to the cortex in the depths of the anterior part of the superior temporal sulcus. One difference from earlier studies is that, whereas in inferior temporal cortex two simple stimuli, which were both effective stimuli for a neuron, produced suppression of the response when presented simultaneously, we found that for neurons in the superior temporal sulcus there was very little suppression of response when two effective faces for a neuron present simultaneously (see Fig. 7).

In a previous study we have shown that the translation-invariance effect for a single face stimulus was independent of whether the stimulus was presented on the contralateral or ipsilateral side of the visual field, relative to the recording site (Tovee and Rolls 1993; Tovee et al. 1993). Consistent with this finding is the result found here that there was no significant difference in the interaction effect described here, depending on whether the peripheral stimulus appeared on the contralateral or ipsilateral side.

References

Ballard DH (1993) Subsymbolic modelling of hand-eye co-ordination. In: Broadbent DE (ed) The simulation of human intelligence. Blackwell, Oxford, pp 71–102

Baylis GC, Rolls ET, Leonard CM (1985) Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. Brain Res 342:91–102

Cowey A, Rolls ET (1975) Human cortical magnification factor and its relation to visual acuity. Exp Brain Res 21:447–454

Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. Brain Research 178:363–380

Desimone R (1991) Face-selective cells in the temporal cortex of monkeys. J Cog Neurosci 3:1–8

Feigenbaum JD, Rolls ET (1991) Allocentric and egocentric spatial information processing in the hippocampal formation of the behaving primate. Psychobiology 19:21–40

Gaffan D, Harrison S and Gaffan EA (1986) Visual identification following inferotemporal ablation in the monkey. Quart J Exp. Psychol 30B:5–30

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the macaque. J Neurophysiol 35:96–111

Gross CG, Mishkin M (1977) The neural basis of stimulus equivalence across retinal translation. In Lateralization in the Nervous System (eds. Harnad S, Doty R, Jaynes J, Goldstein L and G. krauthamer), pp 109–122. New York: Academic Press

Hasselmo ME, Rolls ET, Baylis GC, Nalwa V (1989) Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. Exp Brain Res 75:417–429

Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. Vision Res 20:535–538

Leonard CM, Rolls ET, Wilson FAW Baylis GC (1985) Neurons in the amygdala of the monkey with responses selective for faces. Behav Brain Res 15:159–176

Malsburg C von der (1990) A neural architecture for the representation of scenes. In: McGaugh JL, Weinberger NM, Lynch G (eds) Brain organization and memory: cells, systems and circuits. Oxford University Press, New York, pp 356–372

Merrill EG, Ainsworth A (1972) Glass-coated platinum-plated tungsten microelectrodes. Med Biol Eng 10:662–672

Miller EK, Gochin PM, Gross CG (1993) Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. Brain Res 616:25–29

Mishkin M, Lewis ME, Ungerleider LG (1982) The equivalence of parietopreoccipitial subareas for visuospatial ability in monkeys. Behav Brain Res 6:57–77

Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: Two cortical pathways. Trends Neurosci 6:414–417

Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784

Perrett DI, Rolls ET, Caan W (1982) Visual neurons responsive to faces in the monkey temporal cortex. Exp Brain Res 47:329–342

Richmond BJ, Wurtz RH, Sato T (1983) Visual responses of inferior temporal neurons in the awake rhesus monkey. J Neurophysiol 50:1415–1432

Rolls ET (1984) Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. Human Neurobiol 3:209–222

Rolls ET (1990) A theory of emotion, and its application to understanding the neural basis of emotion. Cogn Emot 4:161–190

Rolls ET (1991) Neural organisation of higher visual functions. Curr Opin Neurobiol 1:274–278

Rolls ET (1992a) Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. Philos Trans R Soc Lond Biol 335:11–21

Rolls ET (1992b) Neurophysiology and functions of the primate amygdala. In: Aggleton JP (ed) The amygdala. Wiley-Liss, New York, pp 143–165

Rolls ET (1992c) The processing of face information in the primate temporal lobe. In: Bruce V, Burton M (eds) Processing images of faces. Ablex, Norwood, NJ pp 41–68

Rolls ET (1994a) Brain mechanisms for invariant visual recognition and learning. Behav Proc (in press)

Rolls ET (1994b) Learning mechanisms in the temporal lobe visual cortex. Behav Brain Res (in press)

Rolls ET, Baylis GC (1986) Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. Exp Brain Res 65:38–48

Rolls ET, Cowey A (1970) Topography of the retina and striate cortex and its relationship to visual acuity in rhesus monkeys and squirrel monkeys. Exp Brain Res 10:298–310

Rolls ET, Tovee MJ (1994) Processing speed in the cerebral cortex and the neurophysiology of visual masking. Proc R Soc Lond [B] 257:9–15

Rolls ET, Burton MJ, Mora F (1976) Hypothalamic neuronal responses associated with the sight of food. Brain Res 111:53–66

Rolls ET, Sanghera MK, Roper-Hall A (1979) The latency of activation of neurons in the lateral hypothalamus and substantia innominata during feeding in the monkey. Brain Res 164:121–135

Rolls ET, Tovee MJ, Purcell DG, Stewart AL, Azzopardi P (1994) The responses of neurons in the temporal cortex of primates, and face identification and detection. Exp Brain Res 101:474–484

Rolls ET, Yaxley S, Sienkiewicz ZJ (1990) Gustatory responses of single neurons in the orbitofrontal cortex of the macaque monkey. J Neurophysiol 64:1055–1066

Rolls ET, Hornak J, Wade D, McGrath J (1994) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol, Neurosurg and Psychiat 57:1518–1524

Rolls ET, Tovee MJ, Purcell DG, Stewart AL, Azzopardi P (1994a) The responses of neurons in the temporal cortex of primates, and face identification and detection. Exp Brain Res 101:474–484

Rolls ET, Hornak J, Wade D, McGrath J (1994b) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol Neurosurg Psychiatr 57:1518–1524

420

Sary G, Vogel R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. Science 260:995–997

Sato T (1988) Effects of attention and stimulus interaction on visual responses of inferior temporal neurons in macaque. J Neurophysiol 60:344–364

Sato T (1989) Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. Exp Brain Res 77:23–30

Tovee MJ, Rolls ET (1992) Oscillatory activity is not evident in the primate temporal visual cortex with static stimuli. Neuroreport 3:369–372

Tovee MJ, Rolls ET (1993) The temporal responses of translation invariance neurons in the primate temporal visual cortex. Perception 22:141

Tovee MJ, Rolls ET, Treves A, Bellis RP (1993) Information encoding and the responses of single neurons in the primate temporal visual cortex. J Neurophysiol 70:640–654

Tovee MJ, Rolls ET, Azzopardi P (1994) Translation invariance in the responses of single neurons in the temporal visual cortical areas of the alert macaque. J Neurophysiol 72:1049–1060

Treves A, Rolls ET (1994) A computational analysis of the role of the hippocampus in nemory. Hippocampus 4:1–18

Wallis G, Rolls ET, Foldiak P (1993) Learning invariant responses to the natural transformations of objects. International Joint Conference on Neural Networks 2:1087–1090

Weiskrantz L, Saunders RC (1984) Impairments of visual object transforms in monkeys. Brain 107:1033–1072