Exp Brain Res (1989) 76: 153–164

Experimental Brain Research

The effect of learning on the face selective responses of neurons in the cortex in the superior temporal sulcus of the monkey

E.T. Rolls, G.C. Baylis*, M.E. Hasselmo**, and V. Nalwa***

University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK

Summary. Neurophysiological studies have shown that some neurons in the cortex in the superior temporal sulcus and the inferior temporal gyrus of macaque monkeys respond to faces. These neurons provided a consistently identifiable substrate with which studies of the storage of visual information were performed. To determine whether face responsive neurons change how much they respond to different novel faces as they become familiar, neurons were tested with two experimental designs. In the first experiment, 22 neurons were tested on their responsiveness to the different members of a large set of novel faces as the set was presented repeatedly until the faces became familiar. 6 neurons altered the relative degree to which they responded to the different members of the set between the first two presentations and subsequent presentations. In a control condition, only 1 out of 17 neurons showed a significant response difference between the first two presentations and subsequent presentations when the experiment started with faces which were already familiar to the monkey. In the second experiment, 26 neurons were tested on their responsiveness to the different members of a set of familiar faces before and after the addition of a novel face to the set. 5 neurons altered the relative degree in which they responded to the different members of the set of familiar faces after addition of a novel face. It is suggested that these changes in neuronal responsiveness to different stimuli reflect the setting up of an ensemble encoded representation of face stimuli. This alteration of neuronal responsiveness as novel faces become familiar suggests that face responsive neurons may store information useful in visual recognition.

In addition to this relatively long-term alteration of relative neuronal responsiveness to different stimuli, it was found that a large number of cells showed a higher mean response to the first presentation of a set of novel faces than to subsequent presentations of the faces. However, the response to the first presentation of a set of familiar faces was also higher than to subsequent presentations in that sequence. This pattern indicates a short term recency effect in the response of these neurons to visual stimuli which is similar to that previously reported (Baylis and Rolls 1987).

Key words: Inferior temporal visual cortex – Face – Visual encoding – Learning – Recognition – Memory

Introduction

When a set of novel visual stimuli is shown, it quickly becomes familiar, and often after only one presentation it is recognised later (Standing et al. 1970; Rolls et al. 1982; Bruce 1988). The inferior temporal visual cortex has been implicated in visual learning and memory (Dean 1976). A population of neurons with responses selective for faces has been described in the inferior temporal visual cortex and in the cortex in the superior temporal sulcus (Desimone et al. 1984; Perrett et al. 1982; Baylis et al. 1985, 1987). Many of these neurons have responses which are different to the different members of a set of faces, so that they could provide an ensemble encoded representation of the identity of a face (Baylis et al. 1985; Rolls 1984, 1989a; Hasselmo et al. 1989; Hasselmo et al. 1989). Because these face responsive neurons have such clearly defined and graded responses, and

Present addresses: * Department of Psychology, University of California at San Diego, La Jolla, CA 92037, USA
** Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA
*** Department of Psychology, University of Delhi, Kamala Nehru College, New Delhi 110049, India
Offprint requests to: E.T. Rolls (address see above)

because they are in a part of the brain which may well be involved in storing the representations of visual stimuli, these neurons were used as a model system to investigate the neurophysiological changes which occur in a population of neurons when that population stores new information.

In this study, we investigated whether individual neurons in this region alter the degree to which they respond to different stimuli when the set of stimuli starts as novel and is repeated until it becomes familiar. This might be expected if the individual neurons in the population interact with each other in such a way that between them they provide a good ensemble encoded representation of visual stimuli. The ways in which such learning can happen in neuronal networks of for example the competitive type are starting to become understood (see e.g. Hinton and Anderson 1981; Rolls 1987, 1989a; Kohonen 1984; Rumelhart and McClelland 1986; Amari 1977; Grossberg 1987). However, most of these studies are theoretical or are performed by computer simulation, and there have been few studies of how in the real nervous system neurons might alter in their relative responsiveness to different stimuli in order to provide across a population of neurons a good representation of information. We also investigated whether the responses of neurons to familiar stimuli alter to those stimuli when a novel stimulus is shown, and learned, as might occur in some types of parallel distributed storage models. An alternative possibility for the functions of neurons in this region is that they are not adaptive filters which learn so that they provide an efficient set of categorisers of the input stimuli, but that they are, at least in the mature monkey, a set of relatively fixed, unmodifiable filters, which do not alter their relative responses when new stimuli are learned by the monkey and stored further on in the visual system.

There is evidence that the responses of neurons in the inferior temporal visual cortex are related to relatively short term, recency, memory. Baylis and Rolls (1987) showed that 26 out of 421 neurons in the cortex of the inferior temporal gyrus and superior temporal sulcus responded differently to novel versus familiar images when there were one or no intervening stimuli, but that this difference was abolished with more than two intervening stimuli. Thus the responses of these neurons reflected whether a visual stimulus had been seen recently, in the preceding one or two stimuli. This result was supported by response changes in a delayed match to sample (DMS) task. In a DMS task using color stimuli (Fuster and Jervey 1981, 1982), neurons were found which fired during the delay period. However, none of these studies was designed to investigate longer term memory, although it was noted that many neurons in this region could provide information useful for the performance of longer-term memory tasks (Baylis and Rolls 1986).

Methods

Recordings were made from both hemispheres of five alert macaque monkeys (2 Macaca mulatta and 3 Macaca fascicularis, weights 3.0–6.5 kg) seated in a primate chair. The activity of single neurons was recorded with glass-insulated tungsten microelectrodes (after Merrill and Ainsworth 1972, but without the platinum plating) using techniques that have been described previously (Rolls et al. 1976). 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 Microvax 2 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. To facilitate latency measurements, the cumulative sum distribution was calculated from the sum peristimulus time histogram.

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 reconstructed from the X-ray co-ordinates taken together with serial 50 μ histological sections which showed the reference electrodes and micro-lesions made at the end of some of the microelectrode tracks.

Stimulus presentation

Stimuli were stored in digital form on a computer disk, and displayed on a video monitor (Microvitec) using a video framestore (AED 512). The resolution of these images was 256 wide by 256 high with 256 gray levels. The monitor provided maximum and minimum luminances of 6.0 and 0.13 footlamberts respectively, and was adjusted internally for linearity to within 3% using a photometer. The computer randomized the order of presentation of these 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. This method allowed completely standardized and randomized presentation of quantitatively specified visual stimuli, and allowed image processing techniques to be applied to the stimuli presented. The monitor on which the images were displayed was placed 1 m from the monkey, and the stimuli subtended 12 degrees at the retina.

The monkeys performed a visual discrimination task during the testing to ensure that they looked at the stimuli. If a small circle (subtending approximately 2 degrees), the positive discriminative stimulus (S+), appeared in the middle of the screen the monkeys could lick to obtain a fruit juice reward, and if a square of the same area and luminance, the negative discriminative stimulus (S–), appeared the monkey had to withhold licking in order to avoid aversive hypertonic saline. A 0.5 s signal tone (400 Hz) preceded the presentation of the stimulus, and if the monkey was fixating correctly before the stimulus appeared, he had sufficient time to perform the discrimination and obtain multiple licks of the fruit juice tube in the short (1.0 s) period in which the stimulus was on. This procedure was designed to ensure fixation of the stimuli (Rolls et al. 1979). If any other stimulus appeared (such as a grating, a 3-D object, or a face), then if the

monkey licked he obtained fruit juice (i.e. all stimuli except the square were treated as S+). The order of presentation of the stimuli was randomized. The EOG recordings confirmed that this procedure resulted in consistent fixation of the stimuli. The monkeys did not perform a recognition task, in which for example only familiar stimuli are rewarded (e.g. Rolls et al. 1982), so that there could be no possibility that differential reinforcement and resulting behavioural responses could account for altering neuronal responses to the stimuli when they were shown repeatedly. Nevertheless the repeated exposure to the stimuli used in this experiment would be sufficient to build a memory trace useful for recognition, as shown for example not only by observations in humans (Standing et al. 1970; Bruce 1988), but also by the facts that exposure of the duration used here sets up a recognition memory on the basis of which neurons in the basal forebrain differentiate novel from familiar stimuli (Rolls et al. 1982), and that such neurons are also set up in monkeys with the training used here, that is in monkeys trained in an association memory and not in a recognition memory task (Wilson et al. 1984).

Face stimuli

This study involved the use of a large number of different faces, which were used as both novel and familiar faces in different experiments. Each face was used once only as novel for each monkey, which limited the number of possible runs. The stimuli consisted of 48 faces in all. These included 4 rhesus and 2 human faces from the standard set used to test for face responses (see Baylis et al. 1985, Fig. 2a), as well as 7 cynomologus faces from a further set used to test responses to this species. A further set of 35 novel faces was made up specifically for these experiments. These consisted of photographs of a large range of individual cynomologus monkeys in the cages of the animal colony. 21 different individuals were photographed, and commonly two photographs of each monkey were digitized, though occasionally one or three were used. The different photographs of each individual were chosen for differences in expression or viewing angle. (This number of individuals was sufficient to ensure that images of sufficient novel individuals were always ready in each monkey in which recordings were made. The different views enabled tests of whether the recognition generalised across views, as shown for example in the lower part of Table 3.) The photographic negatives were digitized using a Scandig 3 (Joyce-Loebl Ltd, Gateshead, U.K.) scanning digitizer of photographs, and stored in an image file with a resolution of 256 × 256 × 8 bits, ready for presentation with the AED 512 framestore.

Non-face stimuli

The responses of the cells were tested to a wide range of non-face stimuli, including sine wave gratings, boundary curvature descriptors, complex 2D images, and three dimensional junk objects, as described previously (Baylis et. al. 1985).

Procedure

While 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 3-D non-face stimuli were shown, to determine whether the response of the neuron was selective for faces. The criteria for a face-selective neuron were that the response to the optimal face stimulus should be more than twice as large as to the optimal non-face stimulus, and that this difference should be significant. (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. Further information on and discussion of the extent to which these neurons have selective responses is given by Baylis et al. 1985; and Rolls 1984, 1989a. 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 criteria, then experiments were conducted as follows.

In experiment 1, the responses of the neuron to the standard set of face and non-face stimuli were first measured for 4–8 iterations of the set of stimuli, to ensure that the responses of the neuron were stable. One iteration consisted of a set of trials on each one of which one of the stimuli from the set was shown. The order of presentation of the stimuli was re-randomized for each iteration. Then the standard set of images was replaced with a set of 4–9 novel face images. (None of these face images had ever been seen before. Most were of monkeys which had never been seen before, but some were in some cases different views of monkeys which had been seen before.) This set of novel face images was presented in random sequences for 7–15 iterations of the novel set, and the neuronal response to every presentation was saved automatically by the computer, as described above. The whole procedure was fully automated to ensure that the stimulus presentation was completely standardized and regular, and care was taken to ensure that there was no break in testing between the standard set and the novel set of images. The neuronal response was calculated from the total number of action potentials occurring on each trial in the period 100–600 ms following stimulus onset. This period was chosen because the cells studied typically responded to visual stimuli with latencies just greater than 100 ms. Recordings of fixation usually confirmed that the monkeys fixated the centre of the screen (the position where the S– was shown on randomly selected trials) during this period of firing rate measurement, but trials with poor fixation (with eye movements of more than 5 degrees) were rejected from the analysis. The neuronal responses were then subjected to analyses of variance, which tested whether the magnitude of the neuronal response to each of the stimuli altered after the stimuli had been shown once or a number of times. In a control set of tests, the procedure was similar, except that when another set of face stimuli was introduced, it was a set which was already familiar to the monkey.

In experiment 2, the responses of the neuron to a set of familiar face stimuli were first measured for 4–8 iterations of the set of stimuli, to ensure that the responses of the neuron were stable. (Examples of the stable responses found before new stimuli were introduced are shown in Fig. 3 in the prenovel period.) Then one novel face was introduced into the set of familiar face stimuli, and the set of images was presented in random sequences for 7–15 more iterations, with the neuronal response to every presentation saved automatically by the computer, as described above. The neuronal responses were measured as described for experiment 1, and then subjected to analyses of variance, which tested whether the magnitude of the neuronal response to each of the familiar stimuli altered after the novel face stimulus was introduced into the set.

The familiar stimuli in this experiment consisted of one of three possible sets of stimuli, with 9, 7 or 5 stimuli. The first possible set used was the standard series for testing for face responses, consisting of 4 faces of different rhesus monkeys, 2 human faces, and 3 non-face stimuli (see Baylis et al. 1985). The second possible set consisted of the 7 cynomologus faces also used frequently to search for face responsive neurons. The final set consisted of 2 rhesus, 2 cynomologus, and one human face drawn from these larger sets. Since these stimuli were used to test for face

responsiveness in every neuron, they became highly familiar to the monkey, with at least 20 presentations each every day. The novel stimuli used in this experiment were drawn from the set of 38 additional photographs described above.

Results – experiment 1

It was possible to complete this experiment for 22 different neurons in four different monkeys. The responses of a neuron (DD0982) which showed a significant change in the relative magnitude of its response to the different faces in the set of novel face images between the first two presentations and subsequent presentations are shown in Fig. 1a. For this neuron, the responses to the different face images a–g became different to each other gradually, with the major changes occurring over the first two presentations of the stimuli. Initially, the profile of the responses to the different faces was flat, showing poor selectivity, but on subsequent presentations the responses to stimuli b and e increased, and the response to stimulus c decreased. These changes were tested statistically with a two-way analysis of variance (ANOVA), in which one (condition) factor was image number (a–g), and the other (group) factor was a first vs a second set of trials (see e.g. Bruning and Kintz 1977). Because most of the changes found occurred in the first one or two trials, the first set of trials consisted usually of the data from trials 1 and 2, and the second set of trials of the data from trials 3 through the last iteration. It was found for this cell that the interaction effect in the two way ANOVA was highly significant ($F(6,30) = 3.8$, $P = 0.006$). This indicates that the relative magnitude of the response of the neuron to the different faces in the set altered between the first set of trials (two iterations) and the second set of trials (5 iterations). There was no significant change of mean firing rate response level between the first two presentations and subsequent presentations ($F(1,5) = 0.14$) (i.e. there was no significant group effect). For this neuron, and for all the neurons tested, there was as expected a highly significant difference between the responses of the neuron to the different images shown (the condition effect), indicating that the neuron discriminated well between the faces in the set, and this condition effect is not commented upon further here.

The responses of a neuron (NN0285) which showed a significant change in the magnitude of its response to the different faces in the set of novel face images between the first presentation and subsequent presentations is shown in Fig. 1b. For this neuron, the response to one familiar face included in the set (stimulus a) decreased after the first presentation. At

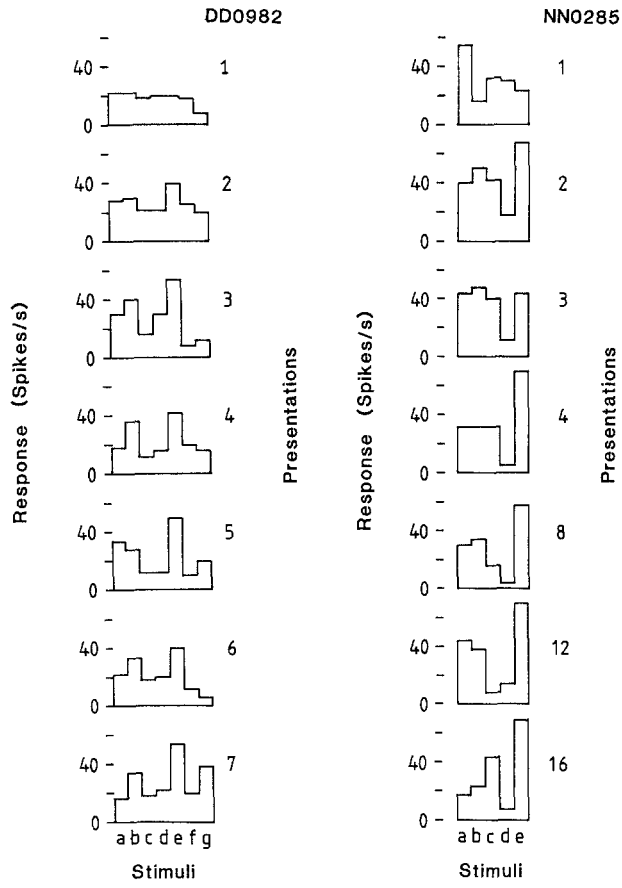


Fig. 1. **a** Firing rate histogram of the response of cell DD0982 showing the effect of 7 iterations of a set of face stimuli. Face stimuli b–g had never been seen before iteration 1 (stimulus a was a familiar face). The relative response of the neuron to the different faces in the set altered, with the greatest alteration occurring in the first two iterations. **b** Firing rate histogram of the response of another cell (NN0285) showing the effect of 16 iterations of a set of novel face stimuli. The relative response of the neuron to the different faces in the set altered, with the greatest alteration occurring in the first one or two iterations

the same time, the responses of the neuron to the novel faces b–e became more differentiated from each other, with the response to faces e and b increasing after the first presentation and the respouse to face d decreasing. These changes were tested statistically with a two-way analysis of variance (ANOVA), in which for this cell the first set of trials consisted of the data from the first presentation of the set of stimuli, and the second set of trials of the data from the remaining 9 iterations of the set of stimuli. It was found for this cell that the interaction effect in the two-way ANOVA was highly significant ($F(4,32) = 6.78$, $P = 0.0003$). This indicates that the relative magnitude of the response of the neuron to the different faces in the set altered between the first set of trials (one iteration of the stimulus set) and the

Table 1. Results from a two-way ANOVA of the first two iterations against subsequent iterations for large sets of novel faces. The F and p values are shown for the interaction between the iteration factor (trials 1 and 2 for group 1, and the remainder of the iterations for group 2) with the image factor in the columns headed Interaction (1, 2). The significance of the iteration factor (i.e. whether there was a change in the mean firing rate between the first and the second group of trials) is shown in the next pair of columns headed iteration. As a control, in the last columns headed Interaction (6, 7), the F and p values are shown for the interaction between the image factor and the iteration factor, with trials 6 and 7 forming group 1 of the iteration factor, and the other iterations forming group 2. No. Pres.: Number of presentations of the set of images

Cell No.	No. Images	No. Pres.	Interaction (1, 2) F	p	Iteration F	p	Interaction (6, 7) F	p
DD0982	9	7	3.80	0.0062	0.14		0.48	0.82
NN0154	9	15	0.62		1.50		0.84	0.52
NN0197	7	16	0.39		4.97	0.045	0.70	0.65
NN0285	5	10	0.76		2.31		1.02	0.41
NN0310	6	15	1.15		3.57		0.31	0.91
NN0318	6	15	0.46		0.23		0.34	0.89
NN0322	6	15	1.88		2.78		1.08	0.38
NN0793	7	14	0.82		22.18	0.0005	0.88	0.51
NN0794	4	11	2.94	0.02	1.32		0.18	0.91
NN0806	6	13	1.42		7.35	0.027	0.43	0.83
QQ0218	7	10	3.62	0.0049	3.41		0.47	0.83
QQ0238	7	10	1.04		1.88		1.67	0.15
QQ0239	7	12	6.83	0.0001	17.62	0.0007	1.00	0.43
QQ0241	7	10	0.98		0.26		—	—
QQ0269	7	14	1.26		4.80		1.42	0.25
QQ0280	5	15	1.69		2.27		1.82	0.14
QQ0329	5	15	0.31		0.01		0.49	0.74
QQ0622	7	8	1.33		0.21		1.68	0.15
PD0237	5	6	2.15		27.60	0.0063	0.90	0.48
PD0336	5	7	0.30		4.84		0.45	0.77
PD0344	5	10	3.27	0.023	4.04		0.33	0.86
PD0506	7	7	17.77	0.0001	0.33		1.04	0.41

second set of trials (9 iterations). There was no significant change of mean response level between the first presentation and subsequent presentations ($F(1,8) = 0.11$) (i.e. there was no significant group effect). For this neuron, as noted above, there was as expected a highly significant difference between the responses of the neuron to the different images shown (the condition effect), indicating that the neuron discriminated well between the faces in the set.

The results for all the cells analyzed are shown in terms of the effects found in the ANOVAs in Table 1. The first group of trials for these analyses of variance consisted of presentations one and two of the stimuli (i.e. the first two iterations of the novel stimulus set), and the second group of trials consisted of all the remaining trials. A total of 6 out of 22 cells showed a significant interaction between groups and treatments, whereas one would be expected by chance (0.05×22). This difference is highly significant (binomial test for the probability of obtaining 6 results significant beyond the 0.05 level in a group of 22 : $P < 0.001$). One of these cells, and 5 other cells showed significant group effects, i.e. an alteration of mean firing rate between the first and the second groups of trials. This reflected a slightly larger mean level of firing on the first one or two occasions on which a given stimulus was seen. The conditions factor is not listed, but was always highly significant, reflecting the fact that the cells showed selectivity between faces.

As a control against effects due to random variations of the neuronal responses, the same statistical test was applied to two later presentations (presentations 6 and 7) of the same large set of novel image data. The results are listed in the last two columns of Table 1. No cells had a significantly different relative magnitude of responsiveness to the different members of the set of stimuli when this pair of presentations was compared to other presentations, indicating that the change in the relative magnitude of the neuronal responses to the different stimuli was particularly high between the first two presentations and the later presentations.

To test whether this change of response pattern was a result of short-term rather than absolute 'novelty', two-way ANOVAs were performed with data in which the set of faces was not novel at the start of the experiment, but was instead already familiar. The data for this test came from the presentations of the large sets of familiar faces used in experiment 2 before the novel face was introduced

Table 2. Results from a 2-way ANOVA of the first two iterations against subsequent iterations for large sets of familiar faces

Cell No.	No. Images	No. Pres.	Interaction F	Interaction p	Iteration factor F	Iteration factor p
DD125	7	7	2.47	0.046	11.42	0.02
DD141	7	7	1.77	0.13	0.06	0.81
II005	7	7	0.30	0.93	0.27	0.63
II009	7	7	2.07	0.09	0.34	0.58
II027	6	7	0.33	0.88	0.13	0.74
II064	7	7	0.72	0.64	5.90	0.06
II075	7	10	0.46	0.83	3.89	0.08
II076	7	10	1.78	0.16	9.62	0.02
II080	5	8	1.07	0.39	3.92	0.09
II081	7	10	0.86	0.53	0.71	0.42
II082	5	11	0.26	0.90	2.06	0.18
II083	5	11	0.34	0.84	1.22	0.29
II084	5	10	1.55	0.21	0.77	0.41
II087	7	8	0.46	0.83	0.01	0.93
II089	5	9	0.95	0.45	0.64	0.45
II090	5	9	1.07	0.39	0.29	0.60
II091	5	9	1.30	0.29	2.04	0.19

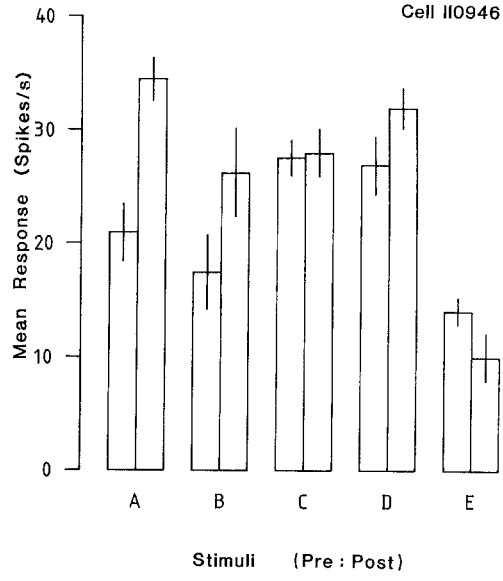


Fig. 2. Example of the change in the relative response of a neuron (II0946) to the different faces in a set of familiar face stimuli produced by the introduction of a novel face. Joined bars for familiar faces A through E represent the mean firing rate response (± sem) to that face before introduction of a novel face to the set (pre, left bar of a pair) and after the novel face had been introduced (post, right bar of a pair)

to the set. (The experimental conditions for this part of experiment 2 were comparable to those of experiment 1 except that the faces were already familiar for the first part of experiment 2). The first two presentations of a set of familiar faces were tested against 6–10 subsequent presentations on 17 cells. (All neurons with seven or more presentations before the novel image was added in experiment 2 were included in the analysis.) The results of this test are shown in Table 2. Only one neuron showed a significant interaction effect in this experiment with familiar faces, and this was the number expected by chance. (Further, this interaction was only just significant, $P = 0.046$.) This difference in the number of neurons with significant interaction effects cannot be due to the number of iterations used in this control experiment, since the cells with significant interactions when novel faces were used as shown in Table 1 still showed significant interaction effects when only the first 7 iterations of their testing were considered.

Because some of the cells showed their major alteration in the relative degree to which they responded to the different visual stimuli between the first and subsequent presentations of a set of novel face images (see e.g. Fig. 1b), the ANOVAs described above were repeated with the responses to the first presentation of each stimulus forming the first group, and the responses during the remaining iterations forming the second group. With these tests, 6 out of the 22 cells showed significant interactions between groups and treatments which again the binomial test shows is highly significantly different from what would be expected by chance: $P < 0.001$. Two of these cells, and 2 other cells showed significant group effects, i.e. an alteration of mean firing rate between the first and the second groups of trials. The conditions factor was always highly significant, reflecting the fact that the cells showed selectivity between faces. With similar statistics applied to the control procedure, it was found that when the responses to presentation 6 were compared to those to the other presentations, no cells had a significantly different pattern of response between presentation 6 and other presentations, indicating that the change in the relative magnitude of the responses to the different stimuli was particularly high between the first presentation and other presentations. Further, in the test in which faces were used which were already familiar because they had been seen on previous testing days, it was found that only one neuron showed a significant interaction effect, and this was the number expected by chance.

It was of interest that altogether 10 of the 22 cells tested with sets of novel face images showed significant interaction effects when either the responses on iteration 1 or the responses on iterations 1 and 2 were compared with the responses on the later iterations. (That is, two of the cells showed significant interac-

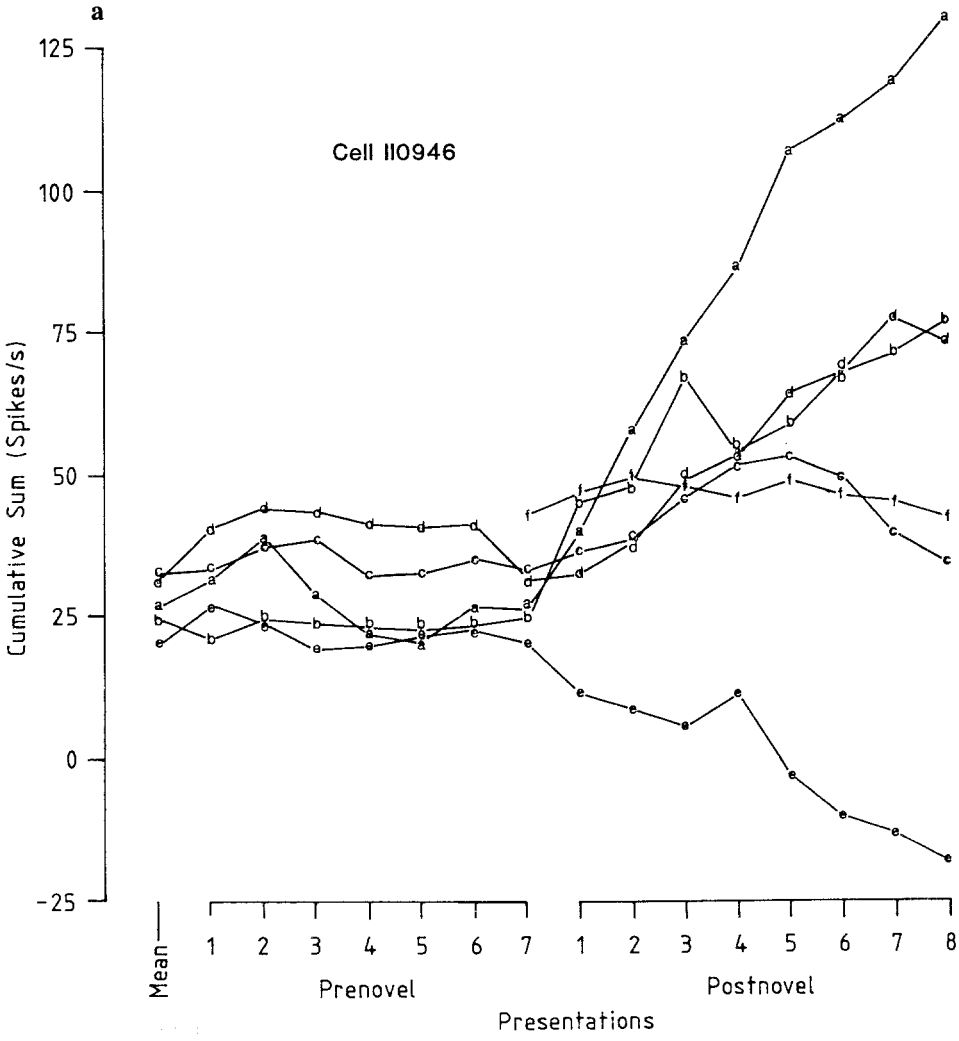
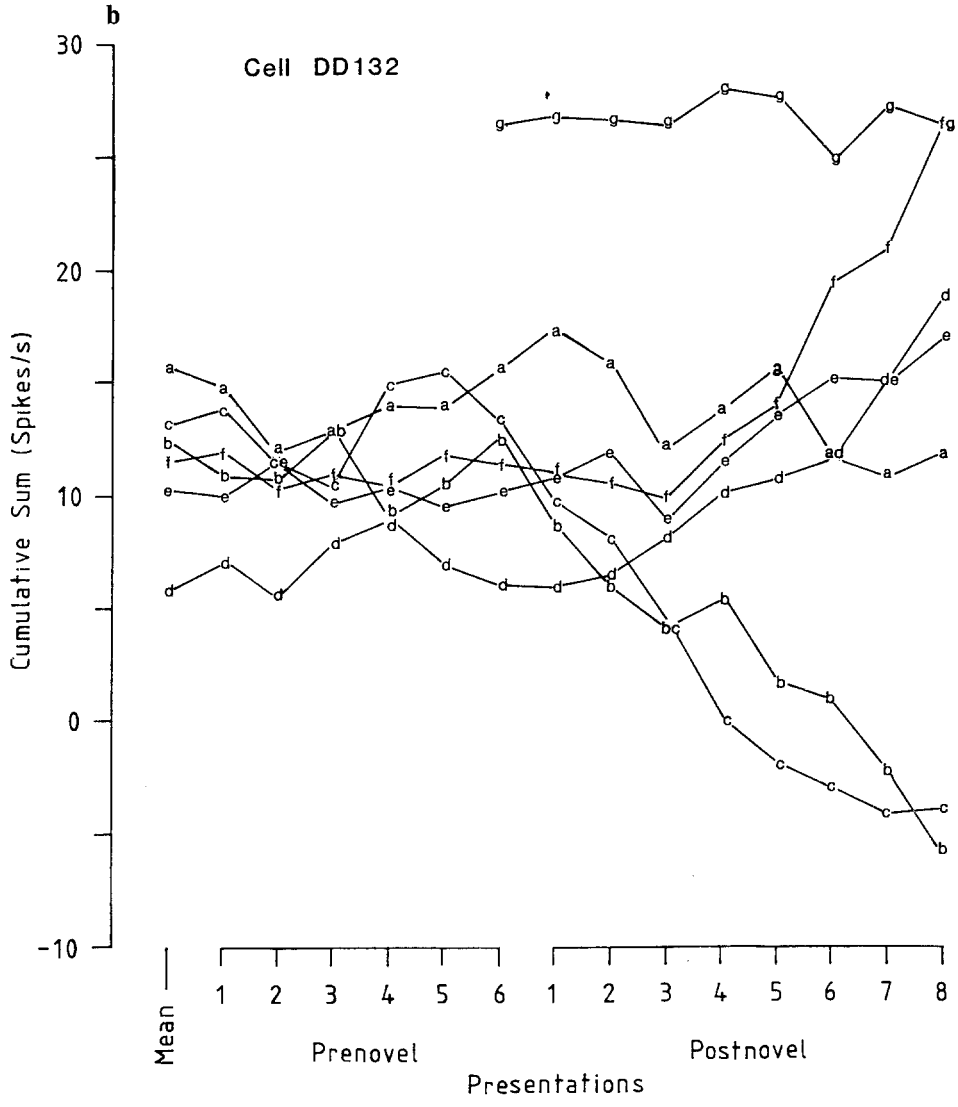


Fig. 3a, b. Examples of the changes in the relative responses of neurons to different familiar faces produced by the introduction of novel faces, with the firing rates represented as cumulative sums before (prenovel) and after the novel face has been introduced (postnovel). The presentation number refers to the iteration number of the set of familiar face stimuli. **a** Familiar faces a through e and novel face f presented to cell II094. **b** Familiar face a through f and novel face g presented while recording from cell DD132

tion effects on both iteration 1 vs the others and on iterations 1 and 2 vs the others.) Thus for the population of 22 neurons analyzed in this experiment, the relative response to the different stimuli altered over either the first one or the first two presentations of the set of novel stimuli for 10 cells. For comparison, only 2 of the 17 cells tested with faces which were already familiar showed a significant interaction effect when either the responses on iteration 1 or the responses on iterations 1 and 2 were compared with the responses on the later iterations, and for both these cells the interaction was only just significant (at the 0.05 level). Thus the majority of the cases in which neurons altered the relative magnitude of their responses to different faces over the first two presentations of the faces occurred when the faces were novel (10/22 cases), not when they were already familiar (2/17 cases). Moreover, the interaction effects found with the sets of novel faces were often very highly significant (see e.g. Table 1), whereas when interaction effects were found with the familiar sets of faces, they were only just significant (see e.g. Table 2).

Results – experiment 2

For this experiment, once a neuron was found to show responses selective for faces, the set of familiar faces was presented in random order, 4–15 times to obtain a suitable baseline level for each familiar stimulus. If the baseline was stable, a novel stimulus was added to the set of faces. This novel stimulus appeared at random within the familiar set, appear-



ing once per iteration of the set, with 6–16 iterations of the set. This method allowed for the investigation of the response of the neuron to the new image as it became familiar, as well as a comparison of the baseline responses to familiar faces before and after the novel stimulus was introduced. It was possible to complete this experiment 26 times in two monkeys (with a total of 23 neurons).

The mean responses to a set of familiar images before the addition of a novel image to the set (prenovel) was compared to the mean response to the same set of images after the addition of a novel image to the set (postnovel). The mean pre- and post-novel responses to five familiar stimuli are shown in Fig. 2 for cell II0946. This cell showed a considerable change in its responsiveness to some of the familiar faces when the novel face was introduced. The cell displayed a decreased response to one image, an increased response to two images, and little change of response to two faces. The change in the relative response to the different stimuli was indicated by the significant interaction in a two-way ANOVA ($F(4,64) = 4.64$, $P = 0.003$) and a significant change of firing rate (group effect) ($F(1,46) = 7.04$, $P = 0.017$).

A more sensitive way of showing small changes in the overall mean in the response of a neuron to an input, and when the change occurs, is the cusum, the cumulative sum of the deviation from mean (Muschaweck and Loevner 1978; Woodward and Goldsmith 1964). The responses of cell II0946 from Fig. 2 are shown in cusum form in Fig. 3a. It is shown in this figure that the neuronal response to face a showed an increase from baseline at postnovel presentation 1 which was maintained at the same high level thereafter, while the neuronal response to other faces showed a slight increase, and the response to face e showed a decrease which was maintained.

Table 3. Results of a two-way ANOVA comparing responses to a set of familiar faces before presentation of a novel image with responses after addition of a novel image to the set. (n = 26)

Cell No.	No. prenovel	No. postnovel	Image No.	Interaction p	Iteration p
DD1250	6	9	726	0.95	0.22
DD1268	4	8	725	0.06	0.04
DD1320	6	8	722	0.02	0.82
DD1411	7	5	731	0.63	0.72
DD1437	6	7	733	0.21	0.001
DD1467	5	6	747	0.004	0.98
II0051	6	10	722	0.25	0.08
II0160	10	8	744	0.13	0.54
II0160	6	8	726	0.10	0.003
II0178	5	9	733	0.22	0.04
II0214	4	7	745	0.90	0.08
II0274	7	9	736	0.97	0.51
II0645	7	6	731	0.95	0.08
II0809	12	7	251	0.09	0.64
II0813	10	9	738	0.050	0.050
II0826	11	10	730	0.008	0.046
II0826	11	8	735	0.07	0.64
II0830	6	7	714	0.66	0.64
II0846	10	8	728	0.74	0.07
II0878	8	9	715	0.14	0.66
II0878	9	10	256	0.11	0.61
II0898	9	10	748	0.53	0.08
II0894	10	16	750	0.40	0.30
II0900	13	10	729	0.84	0.80
II0912	9	11	751	0.31	0.62
II0946	10	8	752	0.0027	0.017

Results of a similar two-way ANOVA with addition of different views of faces which had been seen previously. (n = 9)

II0098	8	8	723	0.11	0.54
II0214	7	8	734	0.54	0.66
II0739	7	8	742	0.30	0.55
II0739	9	6	732	0.70	0.002
II0756	10	8	747	0.15	0.23
II0809	8	11	727	0.53	0.17
II0813	15	12	740	0.09	0.62
II0900	9	11	749	0.16	0.16
II1057	13	9	753	0.20	0.64

With the cusum, the time course of the change can be observed. In Fig. 3a, it is evident that the cusum lines are of generally constant slope past the point of inflection, showing that the change in the response to the stimuli took place relatively rapidly and was maintained relatively steadily thereafter. A further example of a response in cusum form is presented in Fig. 3b for cell DD1320. There was a slightly delayed effect of addition of the novel image, with the response to face f increasing after three postnovel presentations, and the responses to faces b and c decreasing gradually after the addition of novel face g. The interaction effect for pre- vs. post-novel neuronal responses to the stimuli was significant for

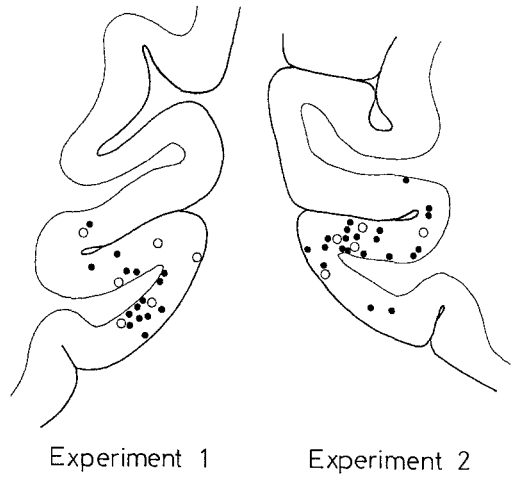


Fig. 4. The locations of neurons recorded in this study; open circles: those showing a significant effect of learning; filled circles: those showing no effect of learning. The cells are represented on coronal sections of the brain taken approximately 6 mm posterior to the sphenoid reference (see Baylis et al. 1987; Aggleton and Passingham 1981) through the superior temporal sulcus, and were recorded within 2 mm of this plane. The sphenoid reference was 18–22 mm anterior to the interaural plane in the different monkeys

this cell ($F(8,96) = 2.37$, $P = 0.02$), and there was no significant alteration of the mean firing rate.

The results of two-way ANOVAs comparing the mean responses across images before and after addition of a novel image to the set are shown in Table 3. The experiment was repeated 26 times (on 23 cells). All of the neurons showed significant response differences to the different faces (the effect of conditions), as expected. A weak criterion for a change in the neuronal response to familiar faces is a significant effect when comparing the responses before and after the novel face was introduced (the group effect), since this could be due to an overall baseline change of the firing rate to the familiar stimuli produced by introducing the novel stimulus. This was found to be significant for 7 cells. This overall effect is significant using the binomial test ($P < 0.001$). A strong criterion for a change is the interaction effect in the ANOVA, which would reflect a change in the relative responses of the neuron to the different familiar faces produced by introducing the novel face. 5 neurons showed significant interaction effects in the two-way ANOVAs, and three of these also showed significant group effects. There are more cases of significant interactions than would be expected by chance (binomial test, $P < 0.01$). Marked interaction effects were not found in 9 different experiments (on 8 neurons, only a preliminary sample) in which a familiar face, but viewed from a different angle to that seen previously,

was introduced into a set of familiar faces, as shown in the lower part of Table 3. The experiment of introducing a novel non-face stimulus was tried on a preliminary sample of four cells, and none of these showed any modification of responsiveness (although one of these cells did alter its responsiveness when a novel face was introduced).

The locations of the cells tested in these experiments are shown in Fig. 4. The cells tested with large sets of novel images in experiment one are presented on the left, while the cells tested with large sets of familiar images with one novel image added in experiment two are presented on the right. (In the left of the figure, two cells were close, and are marked by a single spot.) The cells were located in the cortex in the superior temporal sulcus, or in the cortex forming the inferior temporal gyrus (area TE) (see further Baylis et al. 1987).

Discussion

A significant change in the relative response to different faces between the first one or two presentations and subsequent presentations was found for 10 out of 22 cells tested on a large set of novel faces in Experiment 1 (see Table 1 for a comparison of the responses on the first two iterations compared to the responses on the subsequent iterations). If a later presentation was compared with the other presentations, there were no significant differences in the relative neuronal responses, and only two out of 17 cells showed a significant change in the relative response to different faces when sets of familiar faces were used (cf. Table 2). These findings provide evidence that the relative responses to different faces alter for some neurons in a population of face-responsive neurons as the faces become familiar. It is of interest that this did not occur for all neurons, providing an indication that only some neurons in the population alter their responses when a particular set of new faces is learned. This would lead to the relatively economical use of neurons, many of which would remain available for alteration during the learning of other faces in the future, while at the same time ensuring that many neurons would maintain stable responses even when new faces were learned. Some stability across the population would be desirable so that neurons receiving the output of the neurons analyzed here might continue to respond correctly to previously learned stimuli. The fact that the changes in relative responsiveness found were not large would also tend to assist in such stability. It is thus suggested that the changes in the selective responsiveness of neurons to different stimuli as they are learned reflect tuning of the neuronal responses by experience so that an ensemble of neurons provides a long-term representation of familiar faces.

The results were not due simply to altering patterns of fixation of the stimuli as the faces became familiar, as the following evidence shows. First, the same effects were apparent if only the first 250 ms of the neuronal responses were analysed, before differential eye movements to the stimuli could occur. For example, all the analyses of variance shown in Table 1 columns 4 and 5 were rerun with data collected only in the period 100 to 350 ms after stimulus onset (corresponding because of the typical response latencies of 100 ms to the first 250 ms of neuronal response), and it was found that neurons DD0982, NN0794, QQ0239, PD0344 and PD0506 had significant interactions in the 250 ms response period as well as in the 500 ms response period. (The interaction effect did not reach significance with the small amount of data present in the short response period for neuron QQ0218.) Second, it is being found anyway that the responses of the type of neuron recorded here show little variation in magnitude when the monkey is required fo fixate on different parts of a face (experiments of E.T. Rolls and P. Azzopardi utilising a blink task and search coils to measure the eye position). For these reasons the results described here are very unlikely to be caused simply by different fixation of the stimuli as they become familiar, and it is instead likely that the changing neuronal responses reflect their tuning by experience so that an ensemble of neurons provides a long-term representation of familiar faces.

It is of interest that the major changes in the neuronal responses took place in the first one or two presentations of the faces. Each presentation lasted 1 s. This is the order of time required to build a recognition memory of visual stimuli, in that after a picture has been seen for one or two seconds, it can be recognized later, even after very long periods (see Spoehr and Lehmkuhle 1982; Standing et al. 1970; Rolls et al. 1982). The fact that both the neuronal response changes described here, and the time course of recognition memory, are so similar is consistent with the possibility that the changes in the neuronal responses described here are the neurophysiological substrate of visual recognition memory. Also consistent with this hypothesis are the facts that these neurons are in a part of the visual association cortex which is far removed from early cortical visual processing, and has outputs to multimodal structures such as the amygdala (see Rolls 1989a); and that this cortical area contains some neurons which respond to objects in the environment in object-based, not retinal or viewer-centered co-ordinates (Perrett et al.

1985; Hasselmo et al. 1989), so that it is the type of cortical area where memories of objects may well be stored.

Comparisons of the relative responses to different familiar faces before and after the addition of a novel face to the set showed that in 5 of 26 cases, significant changes in the relative responses of the neuron to the familiar faces occurred. Thus new information encoded by the population of neurons may lead to a change in the relative responses to previously stored patterns. This type of change is what would be expected with a distributed representation, in which each stimulus is encoded by an ensemble of neurons, because the new stimulus would activate some of the neurons activated by familiar stimuli, and the adjustment of synaptic weights for the new stimulus would alter a little the responses to familiar stimuli. This property of distributed representations is made clear by learning in competitive neuronal networks, as described in detail elsewhere (Rolls 1989b, c). In contrast, if information were stored by forming very specific gnostic or grandmother cells (Perrett et al. 1987), then only a very small proportion of neurons would alter their responses when new information was stored, and the probability of finding such a neuron would be very low. The facts that only some of the neurons altered their responses, and that the response changes found were typically small, suggest as above that some stability in the outputs of the population is maintained while at the same time new information is stored in the population.

The ways in which these changes are produced in the neocortex are not yet understood, but the principles of operation of competitive networks are probably relevant to understanding the formation by learning of distributed representations in parallel distributed processing systems (see Rolls 1987, 1989b, c). In such competitive networks, the synapses are Hebb-modifiable, and mutual inhibition between neurons in the population ensures that the neurons remain relatively finely tuned, and that different inputs are stored on different overlapping ensembles of output neurons (see Rolls 1987, 1989b, c; Rumelhart and Zipser 1986; Grossberg 1987). Many of the changes seen in such competitive networks, such as the alteration of the responses of some of the neurons, with the relative responses to some stimuli increasing and to others decreasing, are what is found when competitive neuronal networks are learning and storing new information efficiently (see Rolls 1987; 1989b, c and personal observations of E.T. Rolls).

One possible effect which might have been observed was that the neurons did not differentiate well between the stimuli the first time they were shown, but did learn after a number of presentations to respond differently to the stimuli. Although one cell (DD0982) did show a striking increase of the standard deviation measured across its responses to different faces as they became familiar (see Fig. 1a), most of the neurons did not show this effect. This could be because the neurons had already been adaptively tuned to their present state by previous experience, so that they were already acting as filters, which only required some further tuning to incorporate new information into the network.

The response to novel faces tended to decrease over the first few presentations, as the face became familiar. This is shown by the decrease in the mean firing rate response evident across the first 1–2 iterations of a set of novel faces in experiment 1 (see significant group effects for some cells in Table 1). A similar effect was seen over the first 1–2 iterations during an experimental run with faces which were already familiar in experiment 2 (see significant group effects for some cells in Table 2), so that the effect reflects only the relative recency with which a face has been seen. The duration of the recency effect in this experiment is of the order of 7 trials, in that the set size in Experiment 1 was often 7 faces, so that frequently 7 faces had been seen before a given face was repeated and produced a smaller response. This recency effect is similar to, though lasts a little longer, than that described previously for inferior temporal visual cortex neurons recorded during a serial recognition memory task (Baylis and Rolls 1987). These response characteristics might be ethologically relevant, since social interaction may only require a novelty response at the start of each new encounter, and not during intermittent viewing during such an encounter.

Acknowledgement. This work was supported by the Medical Research Council.

References

Aggleton JP, Passingham RE (1981) Stereotaxic surgery under X-ray guidance in the rhesus monkey, with special reference to the amygdala. Exp Brain Res 44: 271–276

Amari S (1977) Dynamics of pattern formation in lateral inhibition type neural fields. Biol Cybern 27: 77–87

Baylis GC, Rolls ET (1987) Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. Exp Brain Res 65: 614–622

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

Baylis GC, Rolls EG, Leonard CM (1987) Functional subdivisions of temporal lobe neocortex. J Neurosci 7: 330–342

Bruce V (1988) Recognising faces. Erlbaum, London
Bruning JL, Kintz BL (1977) Computational handbook of statistics, 2nd edn. Scott, Foresman, Glenview, Illinois
Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4: 2051–2062
Grossberg S (1987) Competitive learning: from interactive activation to adaptive resonance. Cogn Sci 11: 23–63
Hasselmo ME, Rolls ET, Baylis GC (1989) The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behav Brain Res (in 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 (in press)
Hinton GE, Anderson JA (1981) Parallel models of associative memory. Erlbaum, New Jersey
Kohonen T (1984) Self-organization and associative memory. Springer, Berlin
Merrill EG, Ainsworth A (1972) Glass-coated platinum-plated tungsten microelectrodes. Med Biol Eng 10: 662–672
Perrett DI, Rolls ET, Caan W (1982) Visual neurons responsive to faces in the monkey temporal cortex. Exp Brain Res 47: 329–342
Perrett DI, Mistlin AJ, Chitty AJ (1987) Visual neurones responsive to faces. Trends Neurosci 10: 358–364
Perrett DI, Smith DAJ, Potter DD, Mistlin AJ, Head AS, Milner D, Jeeves MA (1985) Visual cells in the temporal cortex sensitive to face view and gaze direction. Proc R Soc B223: 293–317
Rolls ET (1984) Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. Hum Neurobiol 3: 209–222
Rolls ET (1987) Information representation, processing and storage in the brain: analysis at the single neuron level. In: Changeux J-P, Konishi M (eds) The neural and molecular bases of learning. Wiley, Chichester, pp 503–540
Rolls ET (1989a) Visual information processing in the primate temporal lobe. In: Imbert M (ed) Models of visual perception: from natural to artificial. Oxford University Press, Oxford
Rolls ET (1989b) Functions of neuronal networks in the hippocampus and neocortex in memory. In: Byrne JH, Berry WO (eds) Neural models of plasticity: theoretical and empirical approaches. Academic Press, New York
Rolls ET (1989c) The representation and storage of information in neuronal networks in the primate cerebral cortex and hippocampus. In: Durbin R, Miall C, Mitchison G (eds) The computing neuron. Addison-Wesley, Amsterdam
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, 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 neurones in the lateral hypothalamus and substantia innominata during feeding in the monkey. Brain Res 164: 121–135
Rolls ET, Perrett DI, Caan AW, Wilson FAW (1982) Neuronal responses related to visual recognition. Brain 105: 611–646
Rumelhart DE, McClelland JL (1986) Parallel distributed processing: explorations in the microstructure of cognition. MIT Press, Cambridge, Mass
Rumelhart DE, Zipser D (1986) Feature discovery by competitive learning, Chap 5. In: Rumelhart DE, McClelland JL (eds) Parallel distributed processing: explorations in the microstructure of cognition, Vol 1. Foundations. MIT Press, Cambridge, Mass, pp 151–193
Spoehr KT, Lehmkuhle SW (1982) Visual information processing. Freeman, San Francisco
Standing L, Conezio J, Haber RN (1970) Perception and memory for pictures: single-trial learning of 2560 visual stimuli. Psychon Sci 19: 73–74
Wilson FAW, Rolls ET, Yaxley S, Thorpe SJ, Williams GV, Simpson SJ (1984) Responses of neurons in the basal forebrain of the behaving monkey. Soc Neurosci Abstr 10: 128

Received January 12, 1988 / Accepted December 9, 1988