Exp Brain Res (1982) 47: 329–342

Experimental Brain Research

Visual Neurones Responsive to Faces in the Monkey Temporal Cortex

D.I. Perrett[1], E.T. Rolls, and W. Caan

Dept. of Experimental Psychology, University of Oxford, Oxford, England

Summary. Of 497 single neurones recorded in the cortex in the fundus of the superior temporal sulcus (STS) of three alert rhesus monkeys, a population of at least 48 cells which were selectively responsive to faces had the following response properties: (1) The cells' responses to faces (real or projected, human or rhesus monkey) were two to ten times as large as those to gratings, simple geometrical stimuli or complex 3-D objects. (2) Neuronal responses to faces were excitatory, sustained and were time-locked to the stimulus presentation with a latency of between 80 and 160 ms. (3) The cells were unresponsive to auditory or tactile stimuli and to the sight of arousing or aversive stimuli. (4) The magnitude of the responses of 28 cells tested was relatively constant despite transformations, such as rotation, so that the face was inverted or horizontal, and alterations of colour, size or distance. (5) Rotation to profile substantially reduced the responses of 21 cells (31 tested). (6) Masking out or presenting parts of the face (i.e. eyes, mouth or hair) in isolation revealed that different cells responded to different features or subsets of features. (7) For several cells, responses to the normal organisation of cut-out or line-drawn facial features were significantly larger than to jumbled controls. These findings indicate that explanations in terms of arousal, emotional or motor reactions, simple visual feature sensitivity or receptive fields are insufficient to account for the selective responses to faces and face features observed in this population of STS neurones. It appears that these neurones are part of a system specialised to code for faces or features present in faces, and it is suggested that damage to this system is related to prosopagnosia, or difficulty in face recognition, in man and to the tameness and social disturbances which follow temporal lobe damage and are part of the Klüver-Bucy syndrome in the monkey.

Key words: Neurones – Faces – Monkey – Response properties – Temporal cortex

Introduction

Face pattern processing is a topic of considerable interest particularly since it has been found that very young infants can be differentially responsive to face patterns (Goren et al. 1975; Meltzoff and Moore 1977; Sackett 1966) and because restricted brain damage in human patients can sometimes disturb recognition of faces (prosopagnosia) while leaving the perception of other objects relatively unimpaired (Meadows 1974; Whiteley and Warrington 1977). The brain mechanisms underlying the visual analysis of face patterns and indeed complex patterns in general are, however, unknown.

We were therefore interested when, during investigations of neuronal activity underlying visual discrimination learning (Rolls 1981a), we observed cells within one particular region of the temporal lobe, the cortex in the fundus of the STS, which were responsive to the sight of faces but not to the visual discriminanda or to other stimuli (Perrett et al. 1979; Perrett, unpubl. data). There are a number of ways in which such apparent visual selectivity might arise. Investigations of the basis for the selectivity of the responses of these neurones are described here. These investigations had the aim of advancing our understanding of visual information processing and its disorders which follow brain damage.

1 Present address: The Psychological Laboratory, University of St. Andrews, St. Andrews, Fife, KY 16 9JU, Scotland
Offprint requests to: Dr. D.I. Perrett (see present address)

0014–4819/82/0047/0329/$ 2.80

Methods

Recording Techniques

Techniques that have been previously described (Rolls et al. 1976) were used to record the activity of single neurones with glass-insulated tungsten microelectrodes from the brains of three alert male rhesus monkeys (4.5–6.5 kg b.wt.) seated in a primate chair. The electrical activity of single cells was amplified using conventional techniques (Rolls et al. 1979) and analysed on-line using a PDP11 computer. X-radiographs were used to locate the position of the electrode on each recording track, and the position of cells was reconstructed from histologically verified micro-lesions made at the end of the recordings, and using the X-rays and permanently implanted reference electrodes. Galvanic skin response (GSR) recording was performed with silver/silver chloride surface electrodes which were attached during experiments to the soles of the monkey's feet with collodion glue (SLE, Croydon, UK). Electrode jelly (Neptic, Smith Pharmaceuticals, Welwyn Garden City, UK) was used to ensure good electrical contact. The monkey's behavioural reactions to the sight of the visual stimuli was monitored through a small hole in the side of the chair, and using a video camera with closed circuit TV. Electro-oculogram (EOG) recordings were made using silver/silver chloride electrodes placed on the skin dorsal and lateral to each eye and high gain differential amplification (Tektronix 502A oscilloscope). These methods were sensitive to changes of fixation greater than 2–5 deg of arc. More sensitive recordings with implanted 1-mm ball electrodes, A to D conversion and on-line PDP11 computer processing revealed similar fixation patterns.

Procedure

The monkeys were first trained on a visual discrimination task in which they could lick to obtain a fruit juice reward at the sight of one stimulus (the positive discriminative stimulus, S+), and had to withhold licking to the other stimulus (the negative discriminative stimulus, S–) to avoid aversive hypertonic saline. The discriminative stimuli were presented on trials in pseudo-random order from behind a fast rise time (less than 15 ms), large aperture shutter (Compur Electronic 5FM 6.4-cm aperture) which opened for 1.0 s after a 0.5-s signal tone (700 Hz) provided to allow the monkey to fixate before the shutter opened. After the task had been acquired, the discriminative stimuli were shown on trials interspersed with other test visual stimuli. Lick responses to these test stimuli were normally neither reinforced nor punished but responses to the discriminative stimuli were always reinforced (to the S+ with fruit juice reward and to the S– with saline). In this situation the monkeys were found to attend to and fixate stimuli on all trials even when the discriminative stimuli were used on approximately every tenth trial. Attention was maintained by including more discriminative stimuli if necessary. Fixation of the stimuli for the period in which firing rate measurements were taken, which was usually 0.5 s long commencing 0.1 s after the shutter opened, was confirmed using EOG recordings.

This shutter-controlled situation was used to compare the responses of individual neurones to a number of types of stimuli including simple geometrical stimuli, complex three-dimensional objects, faces, arousing stimuli, faces viewed under a variety of conditions, parts of faces, and arrays of facial features in normal or jumbled configuration. All stimuli were presented against a uniform background (a large white screen). The duration of shutter opening was normally 1.0 s but for particular tests durations between 0.1 and 10 s were used.

Visual Stimuli

Geometrical Stimuli. High contrast square wave gratings, bars, slits and spots of various sizes were produced by placing black stencil material onto the surface of 1 mm thick transparent perspex sheets. These stimuli were placed behind the shutter either singly or stacked in combination to make more complex shapes.

Three-dimensional Objects. Over 1,000 three-dimensional objects were collected. The objects were chosen to differ from one another in size, shape, colour, surface pattern and texture but for convenience of storage the objects were less than 20 cm long. Since these junk objects varied along different visual dimensions, testing neuronal responses to several of them could potentially reveal selectivity for particular visual characteristics. Objects were held either by long forceps or by hand between 2 cm and 1 m behind the shutter or they were placed on the surface of a matt black board tilted towards the monkey.

Face Stimuli. Both real human faces and photographs of faces were used as stimuli. Real faces with neutral expression together with a large toy chimpanzee face were shown through the shutter in the same way as other objects. 35 mm colour slides were prepared of rhesus monkey faces (looking directly at the camera), and of human faces with neutral expression, pictured against a uniform background. Slides were back-projected using a Kodak Carousel projector with intensity adjustable onto ground perspex or drawing paper screens placed at various distances from the shutter (2 cm to 1 m).

Arousing and Aversive Stimuli

Responses to a variety of arousing and aversive stimuli were tested to determine whether arousal itself could account for the neuronal responses which occurred to faces.

Arousing Auditory and Tactile Stimuli. To elicit general arousal, auditory and tactile stimuli were used. Auditory stimuli included various loud noises and human voices made out of sight of the monkey behind a screen at the front of the primate chair. This screen also allowed the monkey's legs to be touched out of sight to test the effects of tactile stimulation.

Visually Aversive Stimuli. Stimuli that the monkey found aversive (as shown by the monkeys' behavioural responses and the GSRs) included an air puffer, feather duster, a large brush and objects looming towards the monkey. Other potentially arousing or interesting visual stimuli included food and reward-related stimuli (used in a visual discrimination task), stimuli that might be taken by the monkey to mean "human" (e.g. hands or lab coats), and model animals (such as a large snake, a spider with dangling legs, and a centipede).

Transformed Views of Faces

Responses to faces were tested under a variety of viewing conditions which systematically transformed the orientation, size and colour of the face image to determine the conditions under which the cells would continue to respond to face patterns.

2D Representations. Colour, black and white photographs, and projections of rhesus monkey and human faces were used.

Colour. The following Kodak Wratten gelatin filters were placed in front of the shutter when real faces were being tested and in

front of the projector when slide stimuli were used: far red no. 92, red no. 29, yellow no. 58, green no. 74, blue no. 47 and neutral density control filters no. 96.

Size. The normal viewing distance measured from the monkey for tests with real human faces was 20 cm. This was increased up to 2 m to provide up to a ten-fold reduction in image size of the face. The distant view included more of the person's body and background.

Rotation and Inversion. Real faces and photographs of faces looking straight ahead were rotated isomorphically from the normal orientation to be either horizontal or inverted.

Profile. The head of the experimenter or that of a toy chimpanzee was turned 90 deg to either side of the monkey so that it was presented in profile. The direction of gaze of the stimulus face was always straight ahead rather than at the monkey.

Parts of Faces

To determine the extent to which responses to parts of the face could account for a cell's responses to the entire face, parts of the face were blanked out or left in isolation. Real and projected faces were presented in the normal manner but screens were set up between the shutter and face stimuli to restrict the monkey's view of particular parts of the face. For back-projected images transparent perspex sheets with matt black paint or Letraset stencil material were used to form screens of a 1-cm bar, 1-cm slit, 2-cm square hole and a semicircular hole of 6-cm diameter. For real faces, both these screens and white card cut out to form screens of comparable dimensions were used. Responses to the entire face were compared to those to the eye presented singly or as a pair, to the hair (usually the top of the head), or to the mouth region with the mouth closed or slightly open to make the teeth visible.

Structural Configuration of Faces

Line-drawn Face. A composite line-drawn face was constructed with individual facial features drawn on transparent square sheets. Individual sheets contained the following features: a triangle of three dots for a nose; two dots for nostrils; two circles for eyes; two dots for iris/pupils; a large circle for the face outline; a fine vertical grating for hair, extending lower at the edges of the face from a fringe over the centre; a row of vertical lines bisected horizontally for teeth and banded by two symmetrical arcs which formed the mouth outline. Cell responses were tested to the line-drawn face, line-drawn features and to jumbled lines. Responsiveness to individual parts of the face or subregions was assessed with different sheets presented individually. To create a jumbled organization of facial features, the individual square sheets were randomly rotated by 90, 180 or 270 deg.

Face Photographs. Entire local facial features, the nose, mouth, hair and each eye were cut out intact from black and white face photographs. These were then attached to plain grey paper cut out to a shape equivalent to the whole face. Features were attached to correct or incorrect localities to form normal or jumbled arrays for different trials. In all arrays the hair was in its normal upright position.

Treatment of Results

For each cell measures of responses were 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 during this period of firing rate measurement, but trials with poor fixation were rejected from the analysis.

For experiments conducted on many cells, the general results are given together with examples of how individual cells responded. When these examples are in the form of a histogram, the mean firing rate and the standard error of this mean response based on typically four to ten presentations of the stimulus are shown. For experiments conducted on a few cells only, statistical treatment of the results for individual cells is included.

Results

Recording Sites

Small electrolytic lesions placed at the recording site of five face selective units were histologically confirmed to be in the cortex in the fundus of the superior temporal sulcus (STS). In the left and right hemispheres of three monkeys (5.5–6 kg b.wt.), X-ray reconstructions confirmed 497 neurones to have been recorded in the fundus region of the STS (either in the cortex of the fundus itself or in the cortex of the upper bank within 2–3 mm of the fundus). These recordings were made between 5 and 11 mm anterior to the inter-aural plane. Figure 1 (left) gives a lateral view of the rhesus monkey brain with the anterior-posterior spread of recordings stippled along the sulcus. Figure 1 (right) is a reconstruction of a coronal brain section made 8 mm anterior to the inter-aural plane and indicates the position of 18 cells showing selective responses to faces recorded in this monkey between 7 and 9 mm anterior to the inter-aural plane.

Of the 497 cells recorded in the STS region there was a sub-population of at least 48 cells which gave responses to the sight of faces that were two to ten times as large as the responses to other stimuli tested. These cells were classified as face selective as a result of the tests described below. A further 49 cells were responsive to faces but insufficient data were obtained to be certain of the basis of their selectivity. Of 207 other cells that were visually responsive, many (98) preferred moving stimuli, with looming stimuli being particularly effective for some neurones. Some other neurones responded both to faces and to other stimuli, which often included arousing or aversive visual stimuli. Of these different cell types recorded in the STS, only results for the 48 face-selective cells will be considered further here. The distribution of face-selective cells isolated so far is centred on a region of cortex which is close to the junction of area TA with area TE of von Bonin and

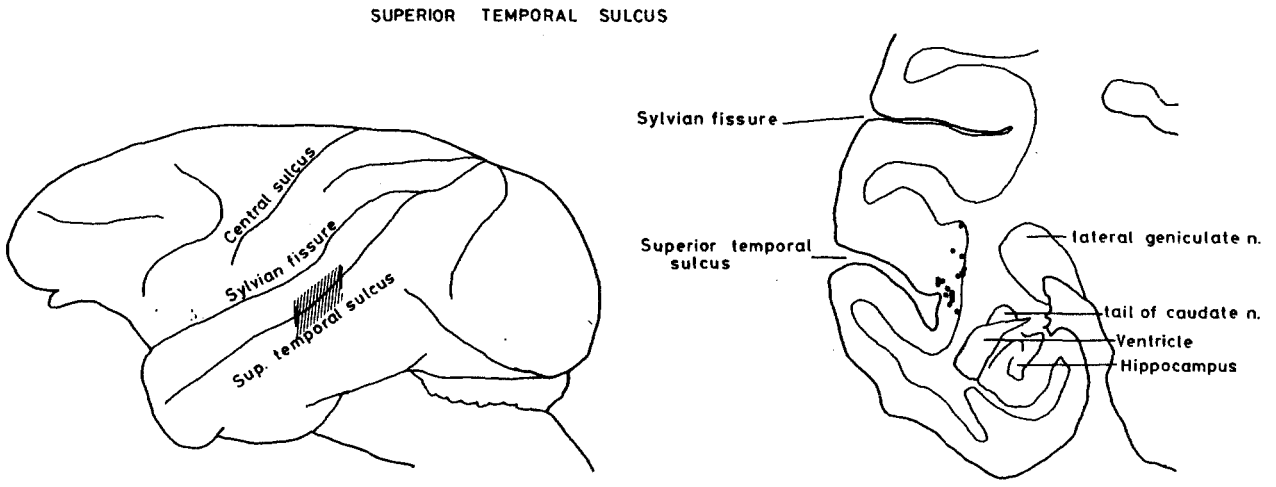

Fig. 1. Anatomical location of cells responding to faces. *Left* Outline drawing of lateral view of left hemisphere of rhesus monkey brain. The stippling along the superior (sup) temporal sulcus indicates the range in anterior/posterior position (5–11 mm anterior to the interaural plane) over which cells responsive to the sight of faces were recorded. *Right* Outline drawing of brain structures evident in a coronal section of the brain of one monkey taken 8 mm anterior to the interaural plane. The positions of 18 cells recorded in this monkey within 1 mm of the plane of the section are reconstructed

Bailey (1947) in the fundus of the STS (see Fig. 1). This region of cortex may be part of what Burton and Jones (1976) have named area T3, and Bruce et al. (1981) have named the superior temporal polysensory area. Neurones responsive to faces were found mainly (but not exclusively) in the left hemisphere. Sampling, however, was biased towards this hemisphere, and thus the possibility of an asymmetry has not been established or excluded.

Temporal Response Characteristics of Cells Responding Selectively to Faces

The cells in the fundus of the STS responding selectively to faces were all found to have excitatory discharges. Individual cells had response latencies between 80 and 180 ms, and the latency distribution is shown in Fig. 2. This range of visual latencies was similar to that of other visual neurones in the STS.

Ninety-two per cent (44/48) of the cells selective for faces gave sustained discharges continuing for as long as the monkey looked at the face stimuli, which was often for many seconds. For five cells, responses to brief presentations (100–300 ms) were tested. For these cells, time-locking of the responses to the onset, duration and offset of the stimulation was observed. Thus, when the presentation was terminated after 100 or 300 ms, responses lasted approximately 100 or 300 ms, respectively (see Fig. 3).

For the majority of cells, responsiveness to real faces was a stable and repeatable finding throughout testing periods lasting often over 1 h and up to 3 h.

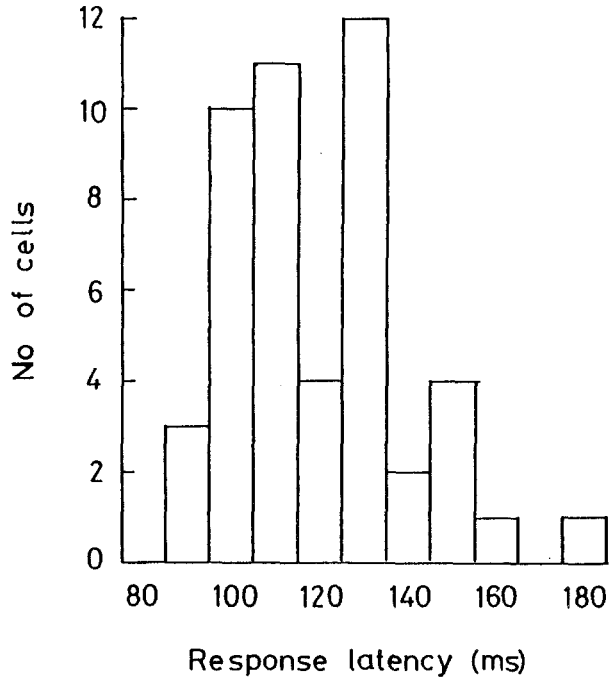

Fig. 2. Response latencies of the neurones in the cortex in the fundus of the STS which responded selectively to faces

Figure 4 is an example of one cell's responses to the face of the experimenter that was shown on trials interspersed amongst trials with other material over a 3-h period. The responses showed some variation in magnitude, but were always present and substantial. When trials on which faces were shown were repeated in quick succession with no alternative stimuli

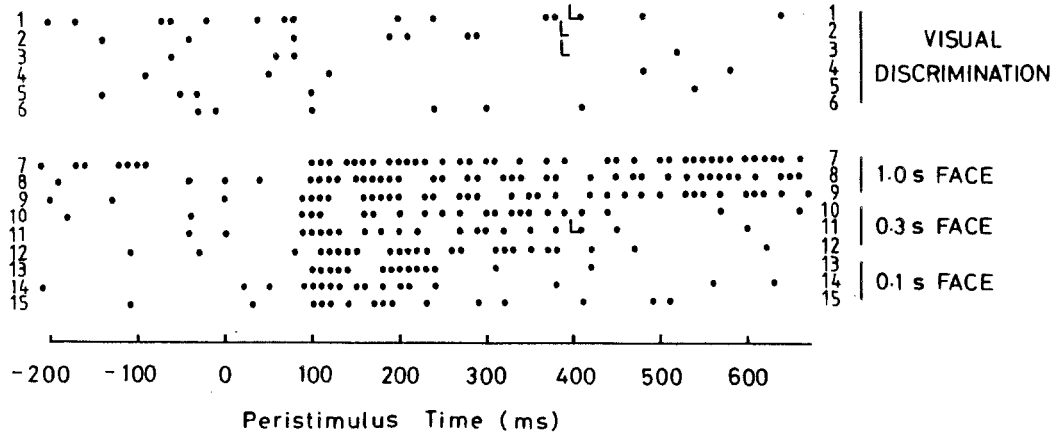

Fig. 3. Visual responses of an STS cell to faces presented for different lengths of time. Each dot represents the occurrence of one or more action potentials in a 10-ms time bin. Each horizontal row of dots represents data from a single trial. Data from different trials which originally occurred in random order have been rearranged for the display. During visual discrimination trials, the monkeys gave a lick (L) response to one of two stimuli, which were presented from behind a shutter at time zero after a 0.5-s signal tone. Time relative to the opening of the shutter is shown along the horizontal axis. The cell recorded was unresponsive to the sight of the discriminative stimuli but gave a large sustained excitatory discharge when the shutter opened to reveal the face of the experimenter

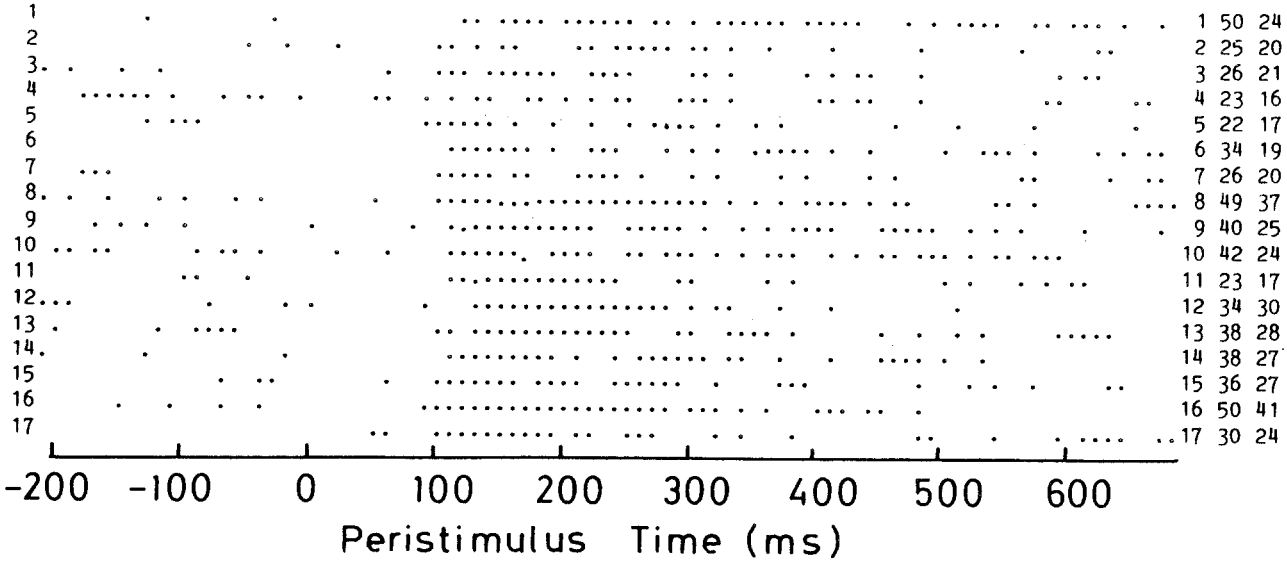

Fig. 4. Consistency of responses to a face over the testing period. Format is as for Fig. 3. Columns at the right side give the total number of spikes occurring on each trial in the time periods 100–300 ms and 100–600 ms after shutter opening. All rows of dots represent activity from trials (over a 3-h period) on which the face of the experimenter was presented

intervening, only five of 23 cells tested showed a tendency for the responsiveness to faces to decline, while the other cells showed no marked tendency for responses to habituate. If responses to stimuli other than faces were present they tended to be transient in nature and often habituated or occurred inconsistently.

Visual Selectivity

Complex Visual Stimuli. Thirty-eight cells were tested, often with large numbers of junk objects. Of these, three gave weak indiscriminate responses to many objects, and 12 cells showed weak or inconsistent responses to particular junk objects. Weak

responses are defined throughout as being less than half of the magnitude of the responses for faces. In most instances, the responses to faces were four to ten times as large as the weak responses to other stimuli.

Simple Geometrical Stimuli. For simple geometrical stimuli, 26 cells were tested and five showed weak responses to some or all of the stimuli.

The results of this testing for one cell are shown in Fig. 5. Here, the mean and standard error of the responses from several trials have been plotted in histogram form. The cell gave a large excitatory discharge to the sight of a face but not to any of the simple stimuli or to the more complex 3D junk objects.

While the above tests can only give a limited indication of visual selectivity it is important to note that testing with junk objects and geometrical stimuli did reveal dependence on particular visual features for other neurones within the STS and in other parts of the temporal lobe such as the inferior temporal visual cortex (Rolls et al. 1977).

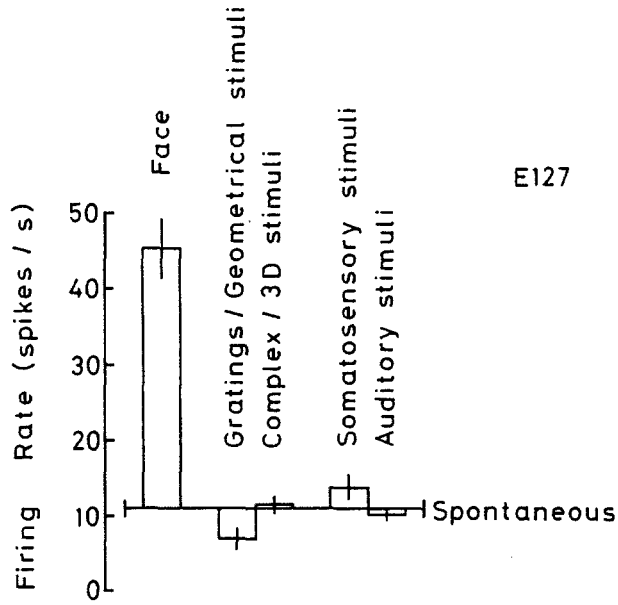

Fig. 5. Visual selectivity. Histogram of the mean (+/– the standard error) of the responses (spikes/s) for one cell are given for different stimuli relative to the cell's spontaneous firing rate. This cell gave large excitatory responses to the sight of faces but did not respond to the sight of simple geometrical stimuli such as bars and square wave gratings or to more complex three-dimensional objects. The cell was unresponsive during somatosensory stimulation (stroking the legs) and to auditory stimuli (loud noises made out of sight)

Effects of Arousal on Neuronal Responses

Auditory and Tactile Stimuli. Tactile stimuli were particularly effective in producing large GSRs, but for 21 cells tested, only five were noted to respond and these cells showed small transient responses. Likewise, among 22 cells tested with arousing sounds only two were found to have any response, and this was transient. The cells are therefore unlike the polysensory STS cells reported by Bruce et al. (1977, 1981). A comparison of the responses of one cell to stimuli in the visual, tactile and auditory modalities is given in Fig. 4. This cell was unresponsive to auditory stimuli even though these included a human voice, and to tactile stimuli, such as touching the legs.

Arousing Visual Stimuli. Of 44 cells tested with the sight of aversive stimuli, eight responded consistently but weakly to more than one stimulus and a further 11 responded inconsistently and weakly to particular aversive stimuli. All the cells that were tested gave stronger responses to faces than to the aversive stimuli despite the greater arousal which could be produced by aversive stimuli, as shown by GSR recordings.

The sight of food and reward-related stimuli (used in a visual discrimination task) produced weak responses from only four of 40 cells tested. Of 39 cells tested with stimuli that might be taken by the monkey to mean "human" (e.g. hands or lab coats), only five showed consistent responses but these were usually weaker than responses to faces. Other potentially arousing stimuli that also proved ineffective in eliciting large responses included model animals and laboratory objects.

Behavioural Reactions. The monkey's facial expression usually remained neutral and often the monkey licked the reward tube in front of his mouth. Various degrees of lip-smacking, grimacing or open mouth threats by the monkey were occasionally observed in response to the sight of faces. These reactions also occurred on some trials with other arousing stimuli. For 12 cells tested, no correlation was observed between the type or intensity of behavioural reaction and the neuronal responses to the sight of faces or other stimuli. It was also found for ten cells tested that making the face rewarding or punishing by making it the positive or negative discriminative stimulus in the visual discrimination had no effect on the magnitude of the responses of the cells to a face.

Figure 6 gives a summary of the responses of one typical cell to the sight of faces and of various types of arousing stimuli. For this cell, although there was a large response to faces, there was only a weak response to aversive stimuli and no response to other

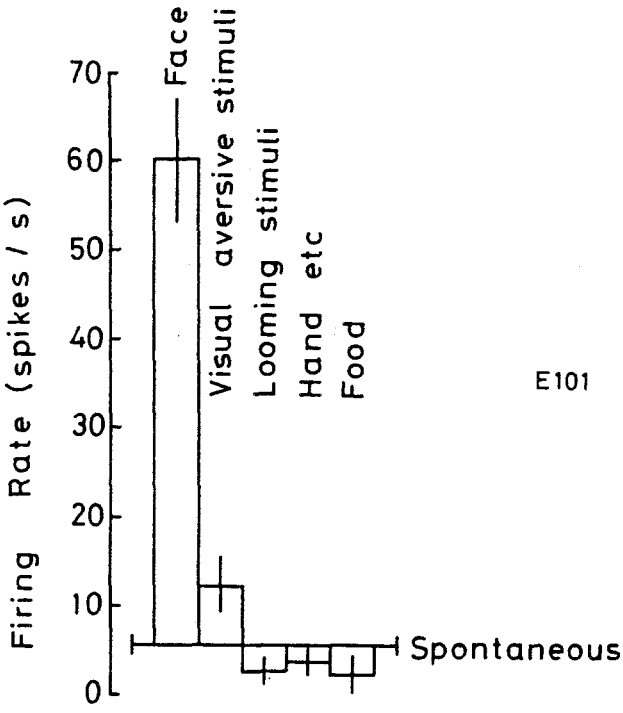

Fig. 6. Arousing stimuli. A comparison of the responses of one cell to the sight of faces and arousing or interesting visual stimuli. The mean and standard error of the responses (spikes/s) are given for the different stimuli relative to the cell's spontaneous firing rate. The cell responded well to the face, but not to the sight of aversive stimuli (such as an air puffer) or to stimuli moving towards the monkey (looming stimuli). The cell also did not respond to stimuli with human meaning (such as a hand) or to the sight of food

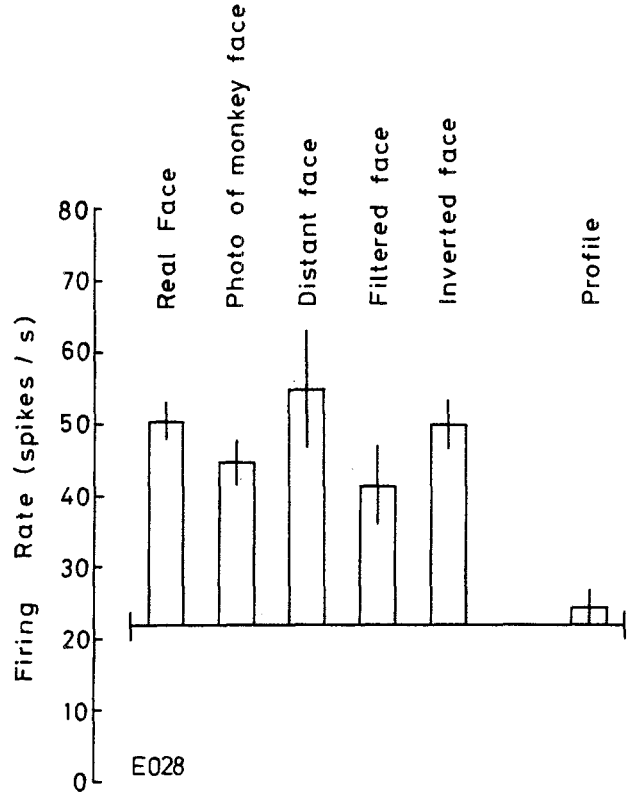

Fig. 7. Transformed faces. The mean and standard error of the firing rate (spike/s) for one STS cell are given relative to spontaneous activity for different views of faces. From top to bottom: experimenter's face at a distance of 2 m, experimenter's face viewed through red, green and blue colour filters, experimenter's face upside down, experimenter's face in profile

arousing stimuli including stimuli looming towards the monkey which produced reliable GSRs and excited many STS movement-sensitive cells. Thus, in general, these tests showed that the STS cells responding selectively to faces did not respond consistently to other arousing stimuli. It seems unlikely therefore that arousal per se can account for the neuronal responses of these STS cells produced by the sight of faces.

Effects of Transformed Views of Faces

The majority of the neurones described here responded to faces which were shown at different distances from the monkey in the laboratory, or were projected onto a screen 20 cm from the monkey. In investigations of the degree to which the responses remained constant despite changes in the viewing conditions, the following results were obtained. The effects of different transformations of faces on the magnitude of the neuronal responses elicited are illustrated in Fig. 7 for one typical cell, and are described for the population below.

2D Representations. For the cell illustrated in Fig. 7, visual responses to a real 3D human face were equivalent in magnitude to the responses to a black and white photograph of a monkey. Seventy-six percent (26/37) of cells tested with 2D pictures of faces were found to respond vigorously to some if not all face pictures tested. A further seven cells (18%) gave significant but weaker responses than those to real faces.

Colour. Since many cells responded to black and white photographs, face colour seemed unimportant. This was confirmed with colour filtered views of faces. For one cell only of 18 tested, there was a slight reduction of response magnitude with colour filtered faces.

Size. Analysis throughout the study with photographs where viewing distance was held constant but

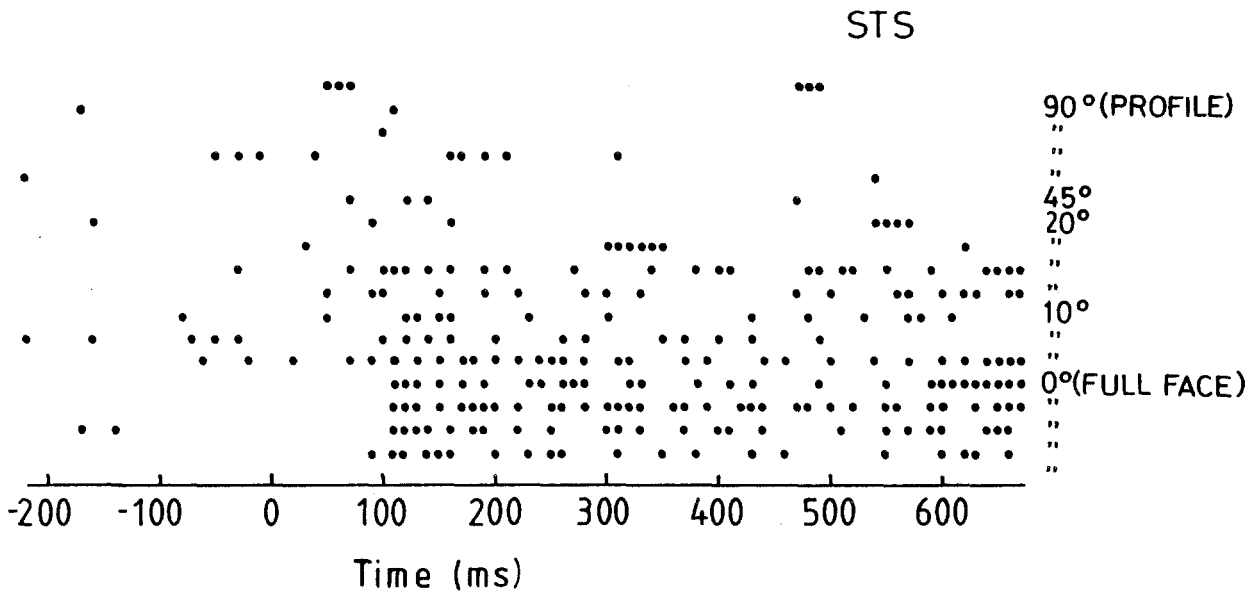

Fig. 8. Responses to face profile. Format as Fig. 3. Activity of one STS cell on individual trials (originally occurring in random order) is displayed on separate lines. For each trial the angle (to the left or right) of the stimulus face toward profile, away from full face, is given in the column on the right

facial size varied from picture to picture and with real faces, failed to reveal obvious effects of size. This constancy of response for different image sizes of faces was confirmed for each of 14 cells tested with real faces viewed at different distances. These distances ranged from 20 cm to more than 2 m.

Orientation. Twenty-one cells tested responded equivalently in magnitude to faces rotated isomorphically, to the horizontal or inverted position.

Profile. One transformation which did affect many cells was the rotation of the head away from the monkey into profile. This reduced or eliminated the responses for 21 of 32 cells. A dramatic example of this effect is illustrated in Fig. 8 where the maximal response was obtained with a full face, and rotation away even by a small angle substantially reduced the neuronal discharge.

Responses to Parts of Faces

In investigations of whether these neurones could respond to elements present in faces, or required the whole face in order to respond, it was found that the majority of cells could continue to respond to the face despite screens obscuring some parts of the face or even half of the face vertically. For individual cells, however, covering up particular regions of the face did substantially reduce or eliminate responses.

Figure 9 illustrates the effects of covering up the eyes or covering the rest of the face for two cells. For the cell on the right side, the presence of the eyes was necessary, and responses were reduced when a bar was placed over them. Consistent with this, presentation of the eyes alone in a slit view elicited a good response from the cell. Thus, for this cell, the eyes were a sufficient and necessary stimulus for the cell to respond. This behaviour is opposite to that of the cell on the left side which responded equivalently to the whole face and to the face with the eyes obscured, and failed to respond when only the eyes were viewed through the slit. Thus, for the cell on the left, the eyes were neither a necessary nor a sufficient part of the face to produce the response found to the whole face. In fact, further tests showed that this cell was responsive to the mouth region and to hair.

Figure 10 illustrates the different behaviour of four cells to face parts. For the cell at the top, a single eye, hair and mouth were ineffective individually in producing responses; the cell did, however, give a good response to the whole face. Cells on the second and third lines responded mainly to one of the facial features tested. The cell on the second line responded quite well to the mouth, whereas the cell on the third line responded more to the eye. The cell on the bottom line responded well to several of the features tested.

Analysis with separate parts of the face (the eyes, mouth and hair) each presented in isolation, revealed for 35 cells that particular parts were effective in

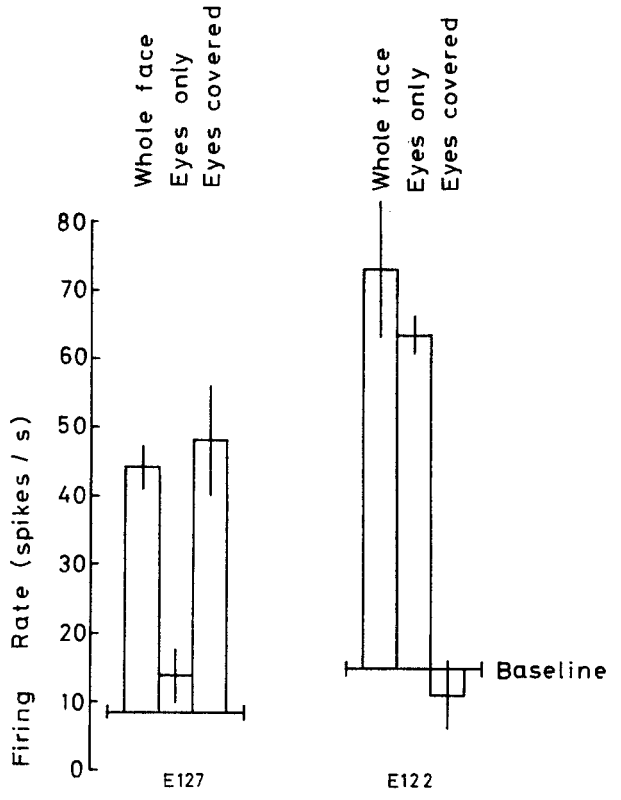

Fig. 9. A comparison of the responses of two STS cells to the presence and absence of eyes in a face. The mean and standard error of the responses (spikes/s) of cells E127 and E122 are given relative to their baseline spontaneous firing rate. For the cell E122 but not for cell E127, eyes were an important part of the face. Cell E122 continued to respond to the eyes viewed alone in a slit but failed to respond to the whole face when the eyes were covered. Cell E127 behaved in the opposite manner, failing to respond to the eyes alone and continuing to respond when the eyes were covered

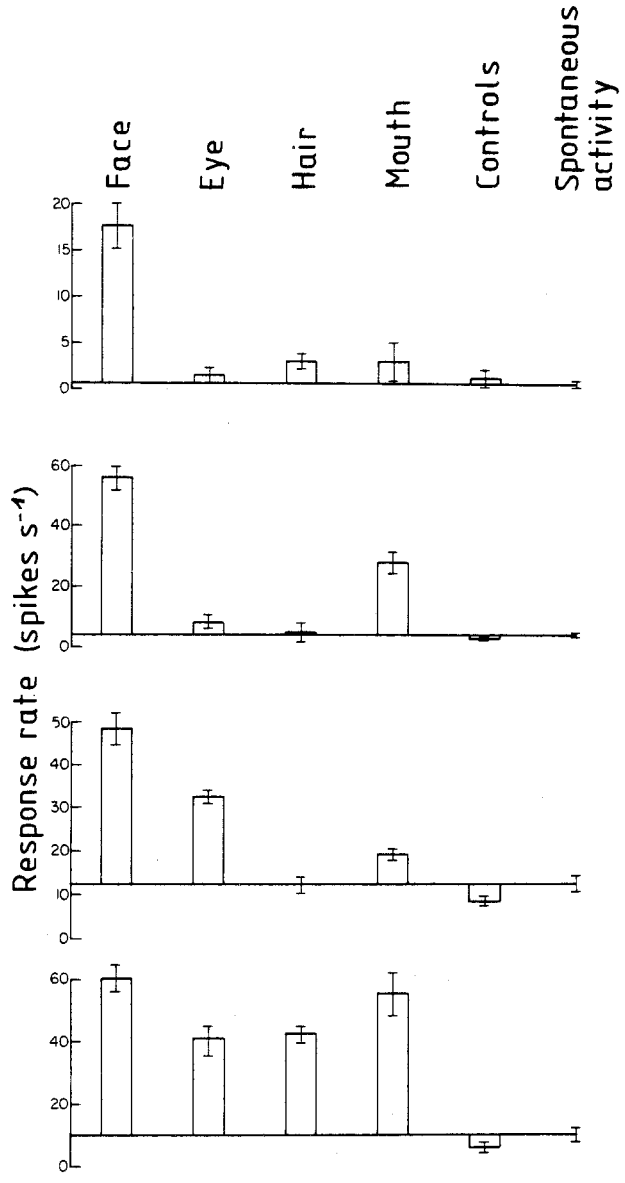

Fig. 10. Responses of cells to the whole face and to individual facial features. The mean and standard error of the responses above the baseline spontaneous activity are given for four cells for the entire face, for facial features tested separately (an eye, hair, the mouth region) and for control visual stimuli (all stimuli tested except faces). For comparison, the scales are drawn so that the response to the whole face is equivalent in size for each cell

producing responses. The eyes were effective in producing substantial or weak but consistent responses for 23 cells (31 tested), the hair for 21 cells (41 tested), and the mouth region for 18 cells (25 tested). The behaviour of cells to different parts of the face was quite hererogenous, cells varying in the number and to which parts they responded. The relative magnitude of the response to effective parts also varied and most cells responded more to the whole face than to ony part presented separately. For seven cells it was found that even the combination of two eyes produced larger responses than one eye presented alone. The differences in the responsiveness to different parts of the face were unrelated to such factors as spontaneous firing rate or the magnitude of the response to the whole face. Thus, cells were found which required all these features to respond, which responded to a face primarily on the basis of one feature, and which responded on the basis of the presence of any one of several different features (see, e.g., Fig. 10).

Structural Configuration of Faces

Line-drawn Faces. Fourteen cells were tested with line drawn faces but most of the cells were not significantly or consistently responsive. For five cells

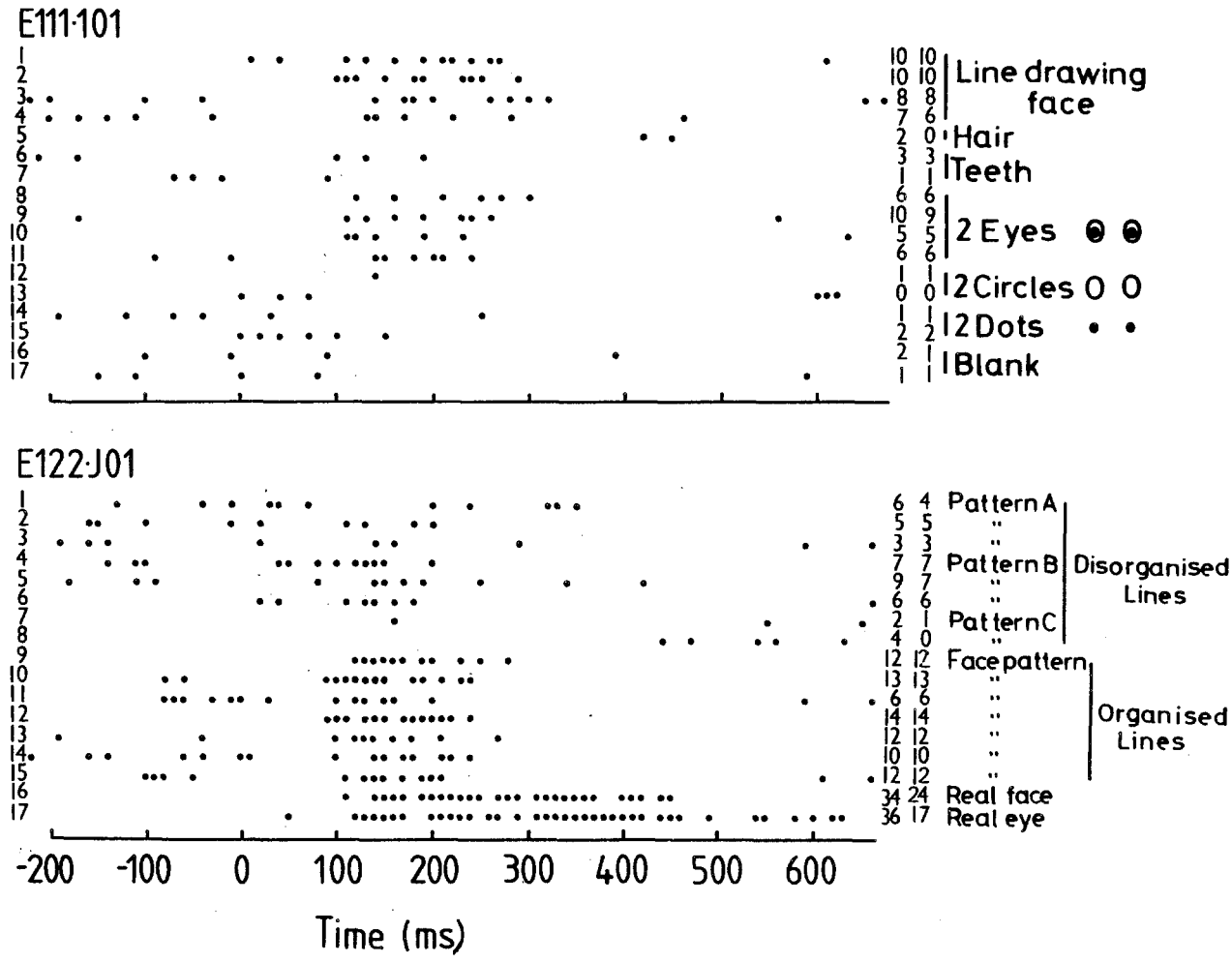

Fig. 11. The effects of the local structure and general organization of line drawn face stimuli. *Upper part.* Format as Fig. 3. The right-hand columns give for each trial the total number of spikes occurring in the 250-ms and 500-ms time periods, commencing 100 ms after shutter opening. The responses of one cell to a line drawn face are compared to the responses of individual line drawn features. This cell responded well to the eye stimuli, a horizontal pair of concentric circles and dots, but failed to respond when the concentric structure was destroyed and two circles or two dots were presented alone. *Lower part.* The responses of a different cell to a line drawn face are compared to the responses to the same lines jumbled into three disorganized patterns (A, B and C). When the lines were organized into a face pattern, responses were significantly larger. The last two lines of the figure illustrate the larger responses of this cell to the sight of a real face or a real eye

weak responses were observed, though it was notable that for these cells, too, real faces were much more effective in producing responses. The results for two cells are illustrated in Fig. 11. Small but consistent responses to the line-drawn face are visible for one cell in the first four lines of the upper figure part, and for the second on lines 9–15 of the lower figure part. In the lower figure part it is also evident that the responses to the face drawings were weaker and more transient than the responses to a real face or even part of the real face, a single eye (lines 16 and 17).

The results for the cell in the upper figure indicate the importance of the local structure of facial features. The response of this cell to the entire line drawn face could be attributed to a large extent to the eye drawings (lines 8–11). When these eye stimuli were further broken down to two circles or two dots, responsiveness disappeared (lines 12–15), indicating the importance of the concentric circular structure.

For the cell in the lower figure, randomly jumbling the composite line features into three different patterns significantly reduced responses. The mean response (for the period 100–350 ms after stimulus onset) for trials with the correctly organized face pattern was significantly larger than the mean response in the equivalent period for the eight trials with the disorganized lines ($t = 5.3$, $df = 13$, $p < 0.001$). This experiment indicates that the structural configuration of the lines of the drawing was an

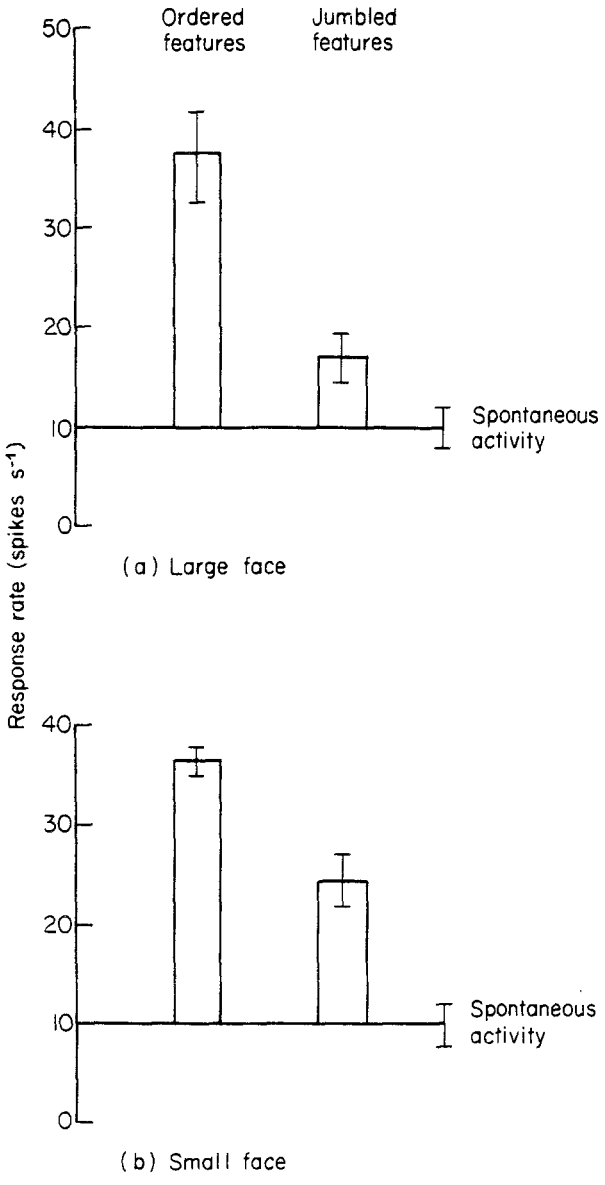

Fig. 12. The mean and standard error of the responses (spikes/s) to the normal and jumbled arrays of face features for two faces. The upper figure gives the mean responses for several different jumbles including jumbles symmetrical about the vertical axis. The lower figure compares responses to normal and one jumbled array for a face one quarter of the size of that used in the upper figure

important factor and that the cell was not simply responding to the line complexity.

Face Feature Photographs. Three of four cells tested were more responsive to the correct organization than to jumbles. The results for one cell are given in Fig. 12. For this cell the responses to two faces are compared to the responses to jumbles of features. For the large face used to obtain the results in the upper part of Fig. 12, several different jumbles (including symmetrical jumbles) were tested and each proved less effective than the correctly ordered face. The results for separate jumbles have been pooled for comparison with results for the properly ordered face. A *t*-test comparison between the responses to normal and jumbled faces was highly significant ($t = 3.98$, $df = 19$, $p < 0.001$). A second smaller face (1/4 of the size) was used to confirm the effects of feature organization. In the lower part of Fig. 12 the response to the correctly ordered features of this face is compared to the response to one jumbled array and again the response to the normal array was larger ($t = 4.03$, $df = 4$, $p < 0.02$).

Responses to Different Individuals

The majority of the neurones described here responded to both human and to monkey faces, and to the faces of different people in the laboratory, whether they were familiar or new to the monkey. Thus, for example, a neurone which responded primarily on the basis of an eye gave large responses to faces of different humans and monkeys, providing that the eyes were visible. In some cases neurones did have different responses to different individuals, and in some cases it was possible to relate this to the degree to which a feature such as hair to which the cell had been shown to respond, was present in particular faces.

Discussion

Possible Explanations of Response Selectivity

There are several lines of experimental evidence which suggest that neither emotional responses, arousal, nor specific motor reactions triggered by the sight of faces can account for the occurrence of the neuronal responses described here.

(1) The time locking of the neuronal responses to the onset and duration of the visual stimulation (see, e.g., Fig. 3) strongly suggests that the cellular discharges were sensory in nature. If the neuronal responses were related to motor or emotional reactions one would expect the latency of discharge onset to be longer and more variable and for the responses to outlast the duration of brief stimuli.

(2) Arousing stimuli generally do not excite the face selective cells in the fundus of the STS despite the larger general autonomic reaction produced by such stimuli than by faces, as shown by GSR recordings. Similarly, there was no observed correspondence between behavioural reactions (facial

expressions, postural changes or eye movements) made on trials with faces or other stimuli and the discharge of face-selective neurones.

(3) Different STS cells responded to different subsets of features of the face. This again suggests that there can be no unitary explanation of the cells' responses in terms of either arousal or emotional reactions produced by the stimuli. Explanations that the responses to faces were due to arousal would predict that all cells would show a similar pattern of reponsiveness for the separate parts of the face, each cell responding most to the most arousing part. Since some cells responded to eyes but not the hair or mouth while others responded most to hair or the mouth, interpretations based on arousal are unlikely to account for the results.

Responses to Transformed Views of Faces

The cells of the STS respond to faces under a wide variety of viewing conditions, and it is therefore unlikely that they are responding to simple attributes of the face, such as colour. Rather, the selectivity in their responses seems based on the presence of complex visual attributes which signify the presence of faces or parts of faces. In contrast to these cells in the fundus of the STS, the cells in other parts of the visual system, including cells in the inferior temporal cortex from which the fundus of the STS receives inputs (Seltzer and Pandya 1978), are in a significant proportion of cases selective in their responses for stimuli of particular orientation, colour and texture (Gross et al. 1972; Rolls et al. 1977). The remarkable degree of tolerance for different images of faces shown by the cells in the fundus of the STS to some extent parallels the constancy of perceptual recognition of objects despite changes in the object's orientation or its viewing distance. There is indeed some evidence for perceptual generalization to different views of faces by monkeys (Rosenfeld and van Hoesen 1979).

It should be noted that while inversion and other visual transformations can impair recognition of the identity of faces by human subjects (Yin 1969), such transformations may not, however, prevent the perception of a stimulus as a face. The particular cells we have described in the fundus of the STS appear to be involved individually in the encoding of the presence of parts of faces or face features, rather than specifying the presence of particular faces. Of course, as a population these cells could convey information specific to particular individuals, and in line with this some cells which did respond differently to different individuals were found. Regardless of this, given their arboreal existence, face recognition in monkeys might be expected to be relatively independent of isomorphic rotation or inversion.

One transformation which did affect many cells was rotation of the head into profile. This reduced or eliminated the response for 60% of the cells tested, and for some, rotation in this plane by even as little as 10 or 20 deg substantially reduced responses. In at least some cases, part of the effect of this transformation was related to the alteration of the view of the eyes. The high sensitivity to head-on faces may reflect the importance of detecting a head-on face, because of the social significance of the face in this orientation. For example, in many primate species including the rhesus monkey directly facing another animal with a maintained stare can act as a threat gesture (Hinde and Rowell 1962).

Responses to Parts of Faces

Tests with parts of the face blanked out or presented in isolation showed that cells of the fundus of the STS responding selectively to faces often responded to component facial features, such as eyes, hair or the mouth region. Since the dependence of responses on different facial features was found while the monkey fixated individually presented features, the response differences cannot be explained by differences in receptive field location. Indeed, although we have not measured the receptive fields, measurements in anaesthetised monkeys indicate that the receptive fields of visual cells in this cortical region are usually large, bilateral and include the fovea (Desimone and Gross 1979; Bruce et al. 1977, 1981).

The cells studied in most cases responded to more than one part of the face, and these parts often appeared to have very different visual characteristics (e.g. the eye and the hair). Further, most cells responded more to combined features of the face than to any of the facial features tested separately. These findings, and the fact that in this particular region of the STS neurones were found which responded to stimuli which happened to have in common that they were parts of faces, suggest that these neurones are involved in visual processing concerned with faces. The grouping of cells in which a particular type of processing is occurring would minimise the length of the interconnecting fibre systems required (Cowey 1979). Further evidence for specialised processing concerned with faces in the STS is that in a study of the inferior temporal visual cortex using the same methods we did not find neurones scattered there which happened to respond selectively to faces or to parts of faces (Rolls et al.

1977). When some neurones were found in a far anterior part of the inferior temporal visual cortex which did resond to faces, they also were grouped together (and were perhaps in a region which projected into the fundus of the STS).

Since visual information does seem to converge upon STS cells we have begun to analyse the effect of combination of and the configuration of face features on their responses. Several cells have been found to give significantly larger responses to the normal organization of cut-out or line-drawn facial features than to control stimuli with scrambled features, including symmetrical jumbles. For seven cells the presentation of two eyes paired horizontally produced larger responses than to one eye alone. These results parallel the findings of increased behavioural responses to normal face configurations and to paired eye patterns in humans and other species (Coss 1968; Scaife 1976; Goren et al. 1975).

General Discussion

The fundus of the STS receives inputs from the inferior temporal cortex and sends efferents to the amygdala, parietal cortex and frontal cortex (Seltzer and Pandya 1978; Aggleton et al. 1980; Jones and Powell 1970; Jacobsen and Trojanowski 1977). It is thus of interest that small, relatively isolated populations of neurones which respond to faces have been found in the anterior inferior temporal cortex (Gross et al. 1972; Rolls et al. 1977), and in the amygdala (Sanghera et al. 1979; Rolls 1981b), parietal cortex (Leinonen and Nyman 1979), and frontal cortex (Pigarev et al. 1979). Bruce et al. (1981) have recently reported seven neurones in the polysensory area in the dorsal (anterior) bank of the STS which responded selectively to faces. It is not clear whether the neurones they recorded were from the same population as those described here in the fundus of the STS. The neurones they described were similar to those described here in that they responded to component features of the face. However, those neurones had very long response latencies (200–300 ms). This could have been due to the use of anaesthetics, or because those neurones were at a later stage of processing to those described here in the fundus of the STS. It is important to note that though our studies were made in a region of the fundus of the STS close to and perhaps part of the superior temporal polysensory area of Bruce et al. (1981), the cells we studied were not polysensory, in that they were generally unresponsive to auditory and tactile stimuli. Both Bruce et al. (1981) and the present authors were impressed by the many different types of response property which could be found in this region, of which the neurones described here are only one set.

The finding of neurones in the temporal lobe with responses which appear to be specialised for faces is of interest in relation to the occurrence clinically of prosopagnosia, a difficulty in identifying individuals from the sight of their face, which is associated with damage to the inferior occipito-temporal region (Meadows 1974; Whiteley and Warrington 1977; Damasio et al. 1982). This type of agnosia, and the neurophysiological findings described here, suggest that there are neural mechanisms specialised for face processing and recognition. Given their responses, the neurons we have described in the fundus of the STS could provide afferent information useful for such a face recognition process. It is also of interest that damage to the amygdala or more widespread temporal lobe regions of the monkey leads to the Klüver-Bucy syndrome (Klüver and Bucy 1939), which includes tameness and a failure to react appropriately to faces (Horel et al. 1975). It may be suggested that this aspect of the Klüver-Bucy syndrome (but not all aspects – see Jones and Powell, 1970; Sanghera et al. 1979; Rolls 1981b) is related to damage to or disconnection of a system in the amygdala which is concerned with emotional responses to faces and which receives face-specific inputs from the STS region in which face processing occurs. Consistent with this suggestion are the findings that the cortex in the fundus of the STS, in which neurones with responses selective for faces were found, projects heavily into the amygdala (Aggleton et al. 1980), and that neurones which responded selectively to faces were found in the amygdala (Sanghera et al. 1979; Rolls 1981b). It is also suggested that damage to this face processing system may contribute to the disruption of social behaviour in the dominance hierarchy which is produced by amygdala damage (Kling and Steklis 1976), and for the normal operation of which recognition of other individuals by the sight of their face may be important. Indeed, it may be that a system specialized for face processing has evolved because of the importance to primates of the rapid and reliable recognition of the faces of other individuals. The neurones responsive to faces in the fundus of the superior temporal sulcus could provide important afferent information to brain systems concerned with identification of and/or social and emotional responses to faces, and it will be of interest to investigate this in the future.

In conclusion, the neurones in the fundus of the STS which we have described here represent a high stage of visual processing related to the analysis of

faces. These responses reflect a considerable synthesis of visual information and demonstrate one, perhaps specialised, form of visual pattern processing where coding for complex patterns is evident at the single cell level.

References

Aggleton JP, Burton MJ, Passingham RE (1980) Cortical and subcortical afferents to the amygdala in the rhesus monkey (Macaca mulatta). Brain Res 190: 347–368

Bonin G von, Bailey P (1947) The neocortex of Macaca mulatta. University of Illinois Press, Urbana

Bruce CJ, Desimone R, Gross CG (1977) Large receptive fields in a polysensory area in the superior temporal sulcus of the macaque. Soc Neurosci Abstr 3: 1756

Bruce CJ, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 46: 369–384

Burton H, Jones EG (1976) The posterior thalamic region and its cortical projection in new world and old world monkeys. J Comp Neurol 168: 249–302

Coss RG (1968) The ethological command in art. Int J Contemp Artist 1: 273–287

Cowey A (1979) Cortical maps and visual perception. J Exp Psychol 37: 1–17

Damasio AR, Damasio H, Van Hoesen GW (1982) Prosopagnosia: Anatomical basis and neurobehavioral mechanism. Neurology Minneap (in press)

Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. Brain Res 178: 393–380

Goren C, Sarty M, Wu P (1975) Visual following and pattern discrimination of face-like stimuli by newborn infants. Pediatrics 56: 544–549

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

Hinde RA, Rowell TE (1962) Comminication by postures and facial expression in the rhesus monkey (Macaca mulatta). Proc Zoo Soc (Lond) 138: 1–21

Horel JA, Keating EG, Misantone LG (1975) Partial Klüver-Bucy syndrome produced by destroying neocortex or amygdala. Brain Res 94: 347–359

Jones EG, Powell TPS (1970) An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain 93: 793–820

Jacobsen S, Trojanowski JQ (1977) Pre-frontal granular cortex of the rhesus monkey. I. Intrahemispheric cortical afferents. Brain Res 132: 209–233

Kling A, Steklis HD (1976) A neural substrate for affiliative behavior in nonhuman primates. Brain Behav Evol 13: 216–238

Klüver H, Bucy PC (1939) Preliminary analysis of functions of the temporal lobes in monkeys. Arch Neurol Psychiatr 42: 979–1000

Leinonen L, Nyman G (1979) Functional properties of cells in antero-lateral part of area 7 associative face area of awake monkeys. Exp Brain Res 34: 321–333

Meadows JC (1974) The anatomical basis of Prosopagnosia. J Neurol Neurosurg Psychiatry 37: 489–501

Meltzoff AN, Moore MK (1977) Imitation of facial and manual gestures by human neonates. Science 198: 75–78

Perrett DI, Rolls ET, Caan W (1979) Temporal lobe cells of the monkey with visual responses selective for faces. Neurosci Lett [Suppl 3], S358

Pigarev IN, Rizzolatti G, Scandolara C (1979) Neurones responding to visual stimuli in the frontal lobe of macaque monkeys. Neurosci Lett 12: 207–212

Rolls ET (1981a) Processing beyond the inferior temporal visual cortex related to feeding, memory, and striatal function. In: Katsuki Y, Norgren R, Sato M. (eds) Brain mechanisms of sensation. Wiley, New York, pp 241–269

Rolls ET (1981b) Responses of amygdaloid neurons in the primate. In: Ben-Ari Y (ed) The amygdaloid complex. Elsevier, Amsterdam, pp 383–393

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

Rolls ET, Judge SJ, Sanghera MK (1977) Activity of neurones in the inferotemporal cortex of the alert monkey. Brain Res 130: 229–238

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

Rosenfeld SA, Van Hoesen GW (1979) Face recognition in the rhesus monkey. Neuropsychologia 17: 503–509

Sackett GP (1966) Monkeys reared in isolation with pictures as visual input: Evidence for innate releasing mechanism. Science 154: 1470–1473

Sanghera MK, Rolls ET, Roper-Hall A (1979) Visual responses of neurons in the dorsolateral amygdala of the alert monkey. Exp Neurol 63: 610–626

Scaife M (1976) Response to eye like shapes by birds. II. The importance of staring, pairedness and shape. Anim Behav 24: 200–206

Seltzer B, Pandya DN (1978) Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. Brain Res 149: 1–24

Whiteley AM, Warrington EK (1977) Prosopagnosia: A clinical, psychological, and anatomical study of three patients. J Neurol Neurosurg Psychiatry 40: 394–430

Yin RK (1969) Looking at upside down faces. J Exp Psychol 81: 141–145

Received November 9, 1981