View-invariant Representations of Familiar Objects by Neurons in the Inferior Temporal Visual Cortex

Michael C.A. Booth and Edmund T. Rolls*

Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

A view-invariant representation of objects in the brain would have many computational advantages. Here we describe a population of single neurons in the temporal visual cortex (IT) that have view-invariant representations of familiar objects. Ten real plastic objects were placed in the monkeys' home cages for a period of time before neurophysiological experiments in which neuronal responses were measured to four views of each object. The macaques performed a visual fixation task, and had never been trained in object discrimination. The majority of the visual neurons recorded were responsive to some views of some objects and/or to the control stimuli, as would be expected from previous studies. However, a small subset of these neurons were responsive to all views of one or more of the objects, providing evidence that these neurons were coding for objects, rather than simply for individual views or visual features within the image. This result was confirmed by information theoretic analyses, which showed that the neurons provided information about which object was being seen, independently of the view. The coding scheme was shown to be sparse distributed, with relatively independent information being provided by the different neurons. Hypotheses about how these view-invariant cells are formed are described.

Introduction

Damage to the temporal lobes of humans and non-human primates can produce impairments in recognizing objects (Weiskrantz and Saunders, 1984; Humphreys and Riddoch, 1987; Farah, 1990). These impairments can occur largely independently of deficits in visual abilities such as acuity, contrast sensitivity and wavelength/colour discrimination. This suggests that the temporal lobe visual cortical areas are concerned with processing visual information at the object level, rather than at the level of simple visual features.

Neurophysiological studies have produced much information on how the temporal cortical visual areas are involved in visual object recognition, revealing that within the inferior temporal cortex (IT) there are neurons that respond to complex visual features (Desimone *et al.*, 1984; Fujita *et al.*, 1992; Tanaka, 1993; Kobatake and Tanaka, 1994; Ito *et al.*, 1995); parts of, or some views of, objects (Mikami, *et al.*, 1994; Nakamura *et al.*, 1994; Logothetis *et al.*, 1995; Logothetis and Sheinberg, 1996); and bodies, faces and hands (Gross *et al.*, 1972; Bruce *et al.*, 1981; Perrett *et al.*, 1982; Desimone *et al.*, 1984; Rolls, 1992; Wachsmuth *et al.*, 1994). Cells that respond to faces have been used as a model system because they are frequently encountered on tracks made into IT. Some of these 'face cells' have responses that are relatively invariant with respect to size and contrast, retinal position, and spatial frequency (Desimone *et al.*, 1984; Rolls and Baylis, 1986; Rolls, 1992; Tovee *et al.*, 1994). Some of these neurons have responses that are relatively view invariant, in that they respond more to some faces than to others, relatively independently of the viewing angle (Hasselmo *et al.*, 1989; Perrett *et al.*, 1991). In these same cortical regions there are also neurons with view-dependent responses. Such neurons might respond, for example, to a view of a profile of a monkey but not to a full-face view of the same monkey (Perrett *et al.*, 1985). These findings of view-dependent and view-independent representations in the same cortical regions are consistent with the hypothesis that view-independent representations are being built in these regions by associating together the outputs of neurons that respond to different views of the same individual (Rolls, 1992). A model of this processing has now been produced (Wallis and Rolls, 1997).

There is less evidence that temporal cortex neurons provide view-invariant representations of objects. Some studies have shown that the responses of some visual neurons in IT do not *depend* on the presence or absence of critical features for maximal activation (cf. Tanaka, 1993, 1996). For example, Nakamura *et al.* (1994) have shown that some IT cells respond to partial views of the same laboratory instrument(s), even when these partial views contain different features. In a different approach, Logothetis *et al.* (1995) have reported that in monkeys extensively trained (over thousands of trials) to treat different views of computer-generated wireframe 'objects' as the same, a small population of neurons in IT did respond to different views of the same wireframe object. An alternative hypothesis, considered here, is that neurons should learn quickly, simply by seeing different views of the same object close together in time – as happens when one inspects and manipulates objects in the real world (for review, see Rolls and Treves, 1998). We note that view-invariant representations of objects by neurons in brain areas such as the IT would be computationally advantageous, in that subsequent processing structures (such as the hippocampus and amygdala) that receive inputs from IT cortex would not need to learn about all possible sizes, positions and views of each object, and there would be easy generalization between one size or position of an object and other sizes, positions or views of the same object (Rolls, 1994, 1995). To investigate whether there are neurons in IT that have similar responses to different views of the same familiar object, we performed the experiments described here. We used real objects, and placed these in the monkeys' home cages, enabling the monkey to pick up the objects and then investigate them from different views. This procedure was used instead of formal visual object discrimination training because it is more like the normal experience that animals (including humans) have with real objects in the world. We then recorded from single IT neurons to determine whether there were any neurons that coded four different views of each object (presented on a video monitor) as being the same. The real objects were toys made of plastic and/or rubber, and were selected to look different from different views, and to contain a limited range of colours. The objects remained in the monkeys' home cages for the duration of the experiment.

Cerebral Cortex Sep 1998;8:510–523; 1047–3211/98/$4.00

One way in which we measured whether neurons coded different views of the same object as being from that one object was by applying information theoretic approaches (for review, see Rolls and Treves, 1998). When the responses of all views of each object were combined, we measured how much information was available in the neuronal responses (on any single trial) about which object had been shown. If the responses to different views of the same object were very different, then little information about which object was presented would be available. If, on the other hand, the neuron responded similarly to different views of the same object, then considerable information about which object was shown should be available on a single trial. We were able to perform this analysis both for single neurons, and for populations of single neurons. Multiple cell information analysis shows how the information increases with the number of cells, providing fundamental evidence about the encoding scheme used. If the information rises in a linear manner with the number of cells, this indicates that the cells convey independent information, and that the code is distributed rather than a local or 'grandmother cell' code (for discussion, see Rolls and Treves, 1998).

Materials and Methods

Recording Techniques

The activity of single neurons was recorded with glass-insulated tungsten microelectrodes in two alert macaque monkeys (*Macaca mulatta* weight ~3 kg; AY, male; BD, female) seated in a primate chair using techniques described previously (Rolls *et al.*, 1990). All procedures, including preparative and subsequent ones, were carried out in accordance with the NIH Guide for the Care and Use of Laboratory Animals, and were licensed under the UK Animals (Scientific Procedures) Act, 1986. The action potentials of single neurons were amplified (Rolls *et al.*, 1979) and converted to digital pulses using the trigger circuitry of an oscilloscope and analysed on-line using an IBM-PC compatible computer. The isolation of single neurons was assured using Brainwave Enhanced Discovery data acquisition, using cluster cutting for off-line data analysis (DataWave Inc.), and establishing that no spikes occurred very close together in time (<3 ms) in the interspike interval histogram. Eye position was monitored and measured with a scleral search coil (Judge *et al.*, 1980) and steady fixation of the visual stimuli was maintained using a 'blink' version of a visual fixation task in which the fixation spot was removed from the video monitor 100 ms before the test visual stimulus was shown. These image presentation procedures have been fully described elsewhere (Tovee *et al.*, 1994).

Stimuli

A different set of 10 objects was used for each of the two monkeys: set A for monkey AY and set B for monkey BD. The two sets of object stimuli are shown in Figure 1. The objects were digitized to the hard disc of a PC using Adobe Capture running under Video for Windows software. Four different views of each object were captured to a Windows bitmap file (512 × 384 pixels), and these files were run together under Adobe Premiere movie-making software to give an AVI video file. (For object set B, the objects were cut out from their original backgrounds and pasted on to a uniform greyscale background.) The 10 objects were then placed in the monkeys' home cages where the monkeys were free to play with them for several weeks before recording started and throughout the duration of the experiment (9 months AY, 6 months BD). During single-neuron recording sessions, the different object views were presented as single frames from the AVI video file. The monitor was placed 50 cm from the monkey, with stimuli subtending ~12° on the retina. The stimuli were presented in conjunction with a computer-generated fixation spot using a Genlock Adaptor video mixing box (Vine Inc.) in the blink fixation task.

Procedure

Tracks were made into the cortex of the superior temporal sulcus (STS) and IT and the responses of isolated neurons were measured to a wide variety of stimuli on the video monitor. These included faces, objects, sine-wave gratings and boundary curvature descriptors (for examples, see Rolls and Tovee, 1995). If the neuron responded to one of the views of one or more of the objects in the home cage to a greater extent than it did to other (non-cage object) stimuli, a series of trials of all 40 stimuli (4 views × 10 objects) was begun. The views were presented in a random sequence generated by the computer presenting 20 views (i.e. 5 objects × 4 views = 20 stimuli) before beginning a sequence again. In some cases the next sequence of 20 stimuli would be from the same five objects, and in others the next 20 stimuli would be from the other five objects. Sequences of trials were continued until each view of each object had been shown to the monkey at least five times.

Once this set of trials had been completed, further experiments were carried out on the cell if it showed view-invariant responses to one or more of the objects in the set, in order to define further the nature of the view-invariance shown. One of these experiments involved presenting the same set of views of the objects in greyscale rather than in colour – for if the cell was truly responding to the object, it might not require a coloured stimulus. In another type of experiment we presented the real three-dimensional object to the monkey, to confirm that the neuronal responses corresponded to those produced by the two-dimensional digitized images. A large-aperture shutter (Compur 5FS, 5 cm diameter) was used to reveal the objects. In further experiments, it was sometimes possible to measure the neuronal responses to views of the cage objects other than the four views contained in the main stimulus set, in order to assess whether the neurons would generalize to any view of a familiar object. However, due to the large number of trials required in the initial stages of testing, these later experiments were not often possible, either because the cell could not be held, or because the monkey became satiated with the fruit juice reward, and no longer worked in the task.

Data Analysis

The firing rates of a cell on each trial were calculated during a 500 ms period starting 100 ms after the onset of the stimulus; neurons in this brain region typically have latencies of 100 ms or more. For each view/stimulus the mean firing rate across the number of presentations was calculated and the standard error computed. The firing rates recorded in response to the 40 different stimuli used in the main experiment were recoded according to object, and a one-way ANOVA of firing rate by object was performed to determine whether the mean firing rates to the 10 different objects (independently of view) were significantly different (at $P < 0.0001$). We note that if a cell does have a significant ANOVA when analysed in this way, this indicates that the firing rates for different objects are different. This in turn implies that the firing rates to the different views of an object tend to be similar, but does not directly address the similarity of the neuronal firing to different views of each object. We therefore also used an information theoretic analysis, as described next, which asks how much information on any single trial (regardless of which view was chosen) is gained about which object was shown.

Information Theoretic Analysis

To show whether the neurons conveyed information about which object was seen, the firing rate responses for the different views of each object were again grouped together to provide a set of firing rates on different trials for each object. We then used an information theoretic approach to determine how much information was available from a single trial from the firing rate measured over a 500 ms period during which an object was shown. If the neuron had different responses to the different views of an object, very little information would be obtained about which object was shown. If, on the other hand, a neuron tended to have similar firing rate responses for the different views of an object, then considerable information about which object was shown would be available from any single trial. The information analysis measured how much information was available (on average) across the set S of 10 different objects, and also the information about each stimulus image s in the set. The information available from each single cell was calculated, and we also calculated the information available from the population of cells. The population measure is of interest because it shows to what extent the cells carry independent information. The procedures are outlined next, and are

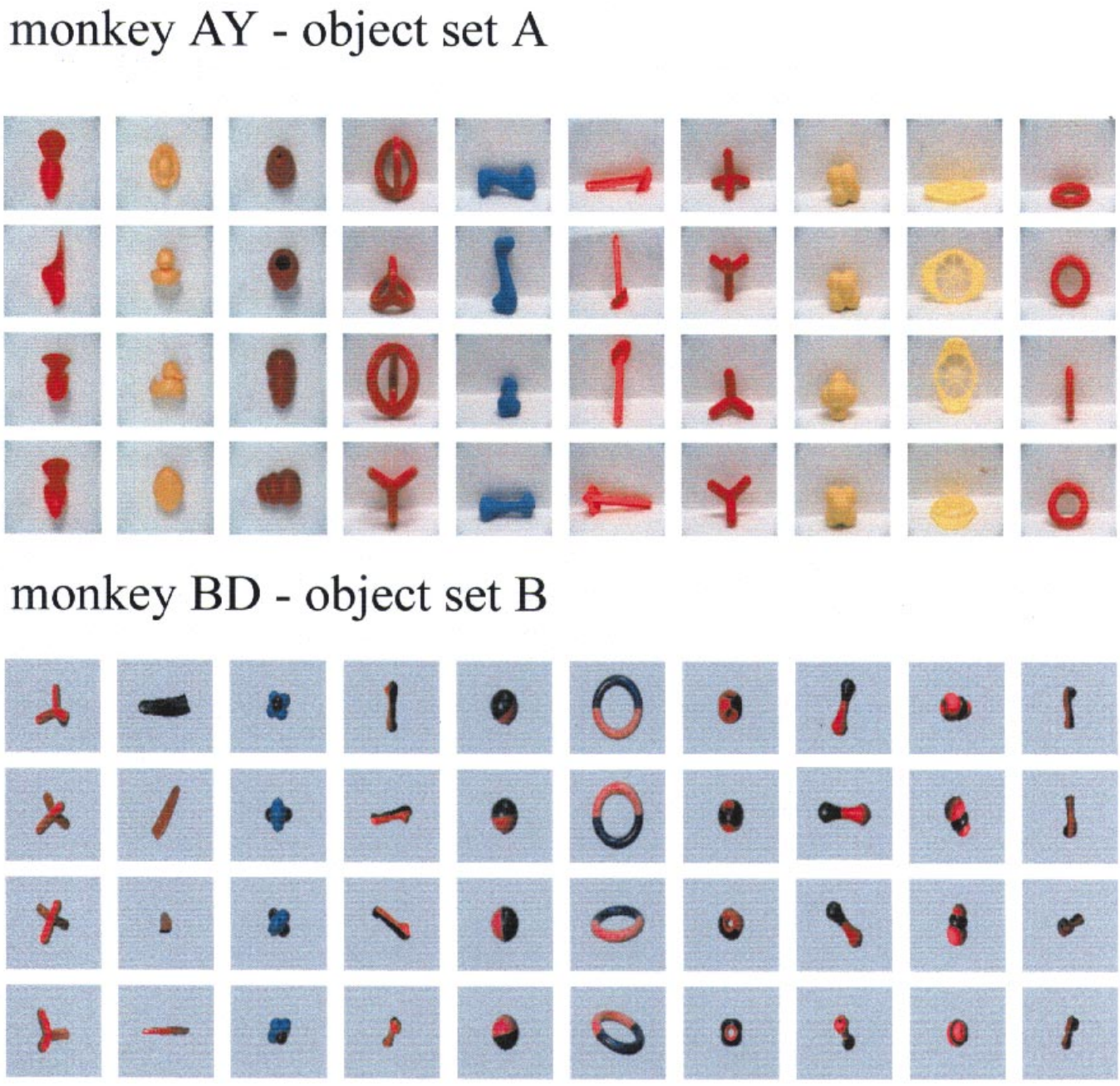

Figure 1. The two sets of object stimuli. The upper set is set A for monkey AY, the lower set is set B for monkey BD

described in more detail by Rolls and Treves (1998, Appendix 2) and by Rolls *et al.* (1997b) for the single cell case, and Rolls *et al.* (1997a) for the multiple cell case.

Single Cell Information Analysis The response r of a neuron to a particular stimulus presentation was computed by measuring the firing rate of the neuron in a fixed time window after the stimulus presentation. Measured in this way, the firing rate responses take discrete rather than continuous values which consist of the number of spikes in the time window on a particular trial, and span a discrete set of responses R across all stimuli and trials. In the experiment, the number of trials available for each stimulus is limited (in the range of 6–15). When calculating the information, the number of values of the firing rate values that can be used must be smaller than (or equal to) the number of trials available for each stimulus to prevent undersampling, and incorrectly high values of calculated information. We therefore quantized the measured firing rates into a smaller number of bins d. We chose here d in a way specific for every cell and every time window, according to the following: d was the number of trials per stimulus (or the number of different rates that actually occurred, if this was lower). This procedure is very effective in minimizing information loss due to over-regularization of responses while controlling effectively for finite sampling biases (Panzeri and Treves, 1996). After this response quantization, the experimental joint stimulus–response probability table $P(s,r)$ was computed from the data [where $P(r)$ and $P(s)$ are the experimental probability of occurrence of responses and of stimuli respectively], and the information $I(S,R)$ transmitted by the neurons averaged across the stimuli was calculated by using the Shannon formula (Shannon, 1948; Cover and Thomas, 1991):

$$I(S,R) = \sum_{s,r} P(s,r) \log_2 \frac{P(s,r)}{P(s)P(r)} \qquad (1)$$

and then subtracting the finite sampling correction of Panzeri and Treves (1996), to obtain estimates unbiased for the limited sampling. This leads to the information available in the firing rates about the stimulus. We did not calculate the additional information available from temporal firing patterns within the spike train, as the additional information is low, often only 10–20%, and reflects mainly the onset latency of the neuronal response (Tovee *et al.*, 1993).

Multiple Cell Information Analysis In estimating the information carried by the responses of several cells, the analysis involved, first of all, constructing pseudosimultaneous population response vectors **r**, occurring, as it were, in what are labelled as 'test' trials [**r** is a vector with one element (or component) for each of the C cells considered]. Each vector consists of the firing rate responses of the C cells to one stimulus *s* on one trial.

From any probability table $P(s,\mathbf{r})$ embodying a relationship between the variable *s* (here, the stimulus) and **r** (here, the response rate vector) one can extract the mutual information

$$I(S,\mathbf{R}) = \sum_{s,\mathbf{r}} P(s,\mathbf{r}) \log_2 \frac{P(s,\mathbf{r})}{P(s)P(\mathbf{r})} \quad (2)$$

where **R** is the response space spanned by the vectors comprising the neuronal activity of all cells. When the probability table has to be estimated as the frequency table of a limited data sample, however, it becomes crucial to evaluate the effects of limited sampling on the information estimate. When **r** is a multidimensional quantity, as it necessarily is if it represents the firing rate of several cells, the minimum number of trials required to sample sufficiently the response space grows exponentially with the dimensionality of that space, i.e. the number of cells considered (Treves and Panzeri, 1995). This rules out, in the case of a limited number of trials, any attempt to evaluate directly the quantity $I(S,\mathbf{R})$. A solution is to use a decoding procedure to estimate for every trial which stimulus was shown. We call this guessed stimulus *s′*. The percentage of correct guesses for different numbers of cells is obtained directly. The information about the true stimulus *s* that was shown can then be measured from the guesses *s′* as follows

$$I(s,s') = \sum_{s} \sum_{s'} P(s,s') \log_2 \frac{P(s,s')}{P(s)P(s')} \quad (3)$$

$P(s,s')$ is the probability, averaged over all the test trials, that a stimulus *s* is shown and is decoded as *s′*. The information value obtained (called I_p by Rolls and Treves, 1998, eq. A2.28) is based on the probabilities that a given stimulus is decoded as each stimulus in the set, and as such is an estimate of equation 2, the direct way to calculate the information. This procedure, the correction for the limited number of trials and the decoding procedure are described in detail by Rolls *et al.* (1997b, 1998). The decoding procedure involved comparison of each response vector to the mean population response vector to each stimulus, as derived from a different set of 'training' or reference data, in order to estimate the relative probabilities $P(s'|\mathbf{r})$ that the response **r** had been elicited by any one stimulus *s′* in the set. Separating the 'test' from the 'training' data is called cross-validation, and was performed using a *jack-knife* technique in which one of the available trials for each stimulus was used for testing, and the remaining trials for training. A probability estimator (PE) algorithm was used which reconstructs the correct Bayesian probabilities from the data assuming a Gaussian distribution of the neuronal responses (Rolls *et al.*, 1997b). An alternative decoding is to measure the likelihood of the stimuli using a measure of the distance between the test vector and each of the average response vectors to each stimulus. The greater the distance, the lower is the probability that the stimulus was *s′*. The distance measure could be based on the Euclidean distance (Rolls *et al.*, 1998) or the dot product. The dot product (DP) algorithm just computes the normalized dot products between the current firing vector **r** on a test trial and each of the mean firing rate response vectors in the training trials for each stimulus *s′*. (The normalized dot product is the dot or inner product of two vectors divided by the product of the length of each vector. The length of each vector is the square root of the sum of the squares.) The highest dot product indicates the most likely stimulus that was presented, and this is taken as the best guess for the percentage correct measures. For the information measures, it is desirable to have a graded set of probabilities for which of the different stimuli was shown, and these were obtained from the dot products as follows. The *S* dot product values were cut at a threshold equal to their own mean plus one standard deviation, and the remaining non-zero ones were normalized to sum to 1.

Recording Sites

X-radiographs were taken at the end of each recording session to determine the position of the microelectrode, relative to bony landmarks and the permanently implanted reference electrodes. At the end of the final tracks, microlesions were made in the areas of cortex in which recordings were made to mark typical recording sites (Feigenbaum and Rolls, 1991). Reconstructions of the tracks were made in serial 50 μm histological sections using the positions of the microlesions and the reference electrodes in the histology, and the corresponding X-ray coordinates and the X-ray coordinates of all recorded cells, to determine the locations of all the cells.

Results

We recorded from 737 neurons in the IT of the two hemispheres of two monkeys. Of the 290 (290/737 = 39%) visual neurons recorded, 57 (20%) responded best to faces, and of the remaining 233 visually responsive neurons, 21 (9%) responded to one or more of the cage-objects in a view-invariant manner. That is to say, they responded to all views of one or more of the objects and not to others, as described below. Of the other visual (non-face selective) neurons, 75/202 (37%) responded to some views of some of the cage objects shown on the monitor or to some views of other three-dimensional objects shown to the monkey; and the other 137 visual neurons had responses that either could not be classified as being selective, or were lost before conclusive testing could be carried out. Therefore, of the visual neurons that could be extensively tested, 57/153 (37%) were face selective, 21/153 (14%) had view-invariant responses to one or more of the cage objects, and 75/153 (49%) had feature based or view-dependent responses to visual stimuli.

Firing rate profiles of examples of neurons which responded similarly to different views of the same object, and had responses which were different across the object set, are shown in Figure 2 for monkey AY, and Figure 3 for monkey BD. In Figure 2 it is shown that neuron ay101b was driven by the four images of the apple corer (object 9), but was not driven by other yellow images, or by images with round or elongated sections. It is notable that the first and fourth view of object 9 are (in terms of shape) different from the second and third view. It is also shown in Figure 2 that neuron ay095b fired to the four views of object 4, but did not respond to even very similar images of other objects (e.g. object 7). A further illustration is given in Figure 4, which shows peristimulus time histograms and rasterplots of neuron bd054b's responses to the four views of object 6, with the object views horizontally adjacent to the appropriate plot. Across the population of 21 reported cells, the maximal evoked firing rate had a mean of 60.4 spikes/s and a SD of 29.6, and the spontaneous rate had a mean value of 10.4 spikes/s with a SD of 10.1.

Further Experimental Results

To obtain evidence on whether the neurons were responding to different views of the same object, and that colour was not a crucial factor in accounting for the cell's responses during the testing, we performed experiments in which the colour was removed from the images being presented. An example of such an experiment is shown in Figure 5 (for neuron ay095b) which shows the neuron's averaged response to the four views of

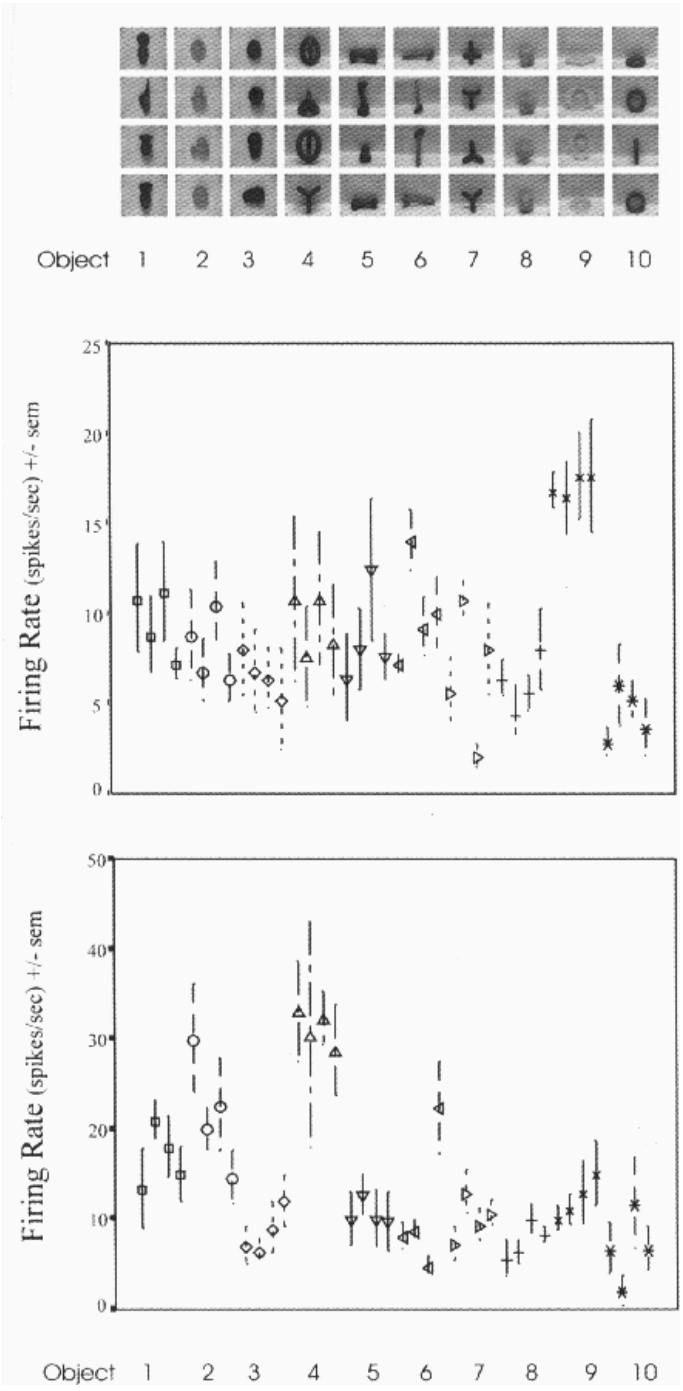

Figure 2. Firing rate profiles of two example neurons (ay101b, upper and ay095b, lower) from monkey AY that responded similarly to different views of the same object, and had responses that were different across objects. The images, grouped by object, run along the abscissa; the firing rates (in spikes/s) are shown on the ordinate. On this and subsequent firing rate profiles, the different symbols and the error bars (solid or broken lines) are solely meant to aid the presentation of the data points as coming from the same object.

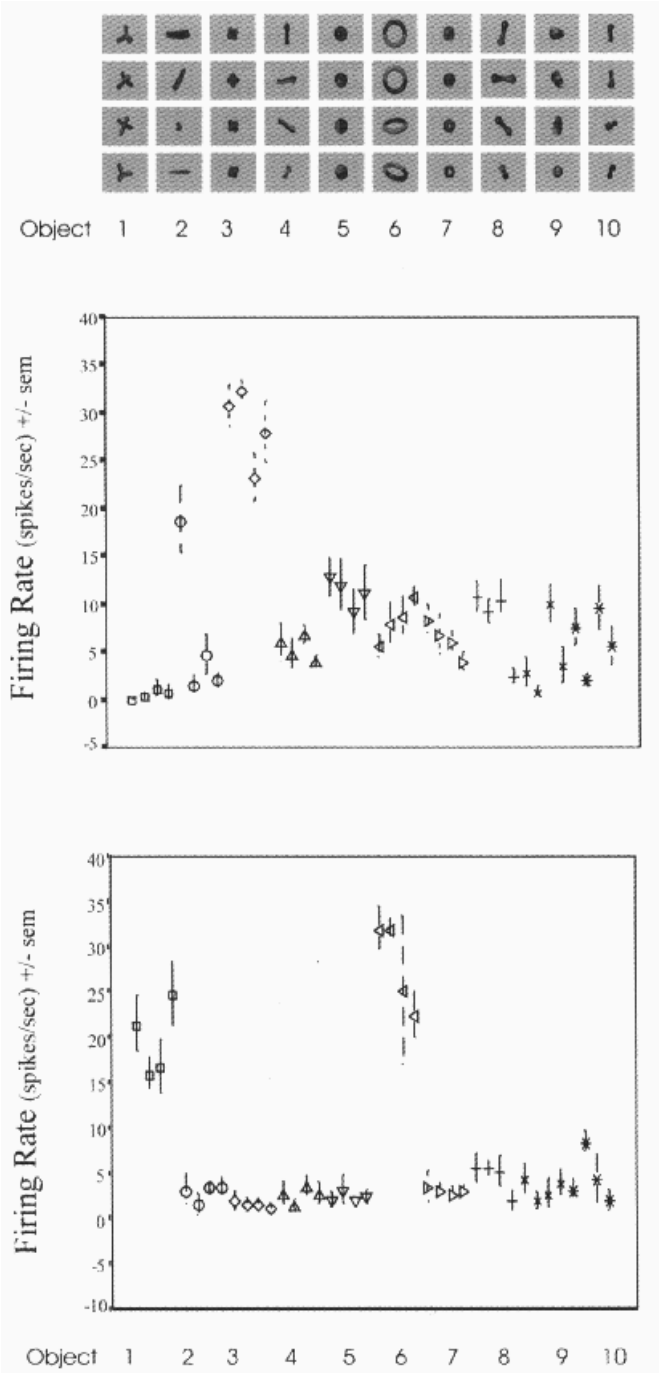

Figure 3. Firing rate profiles of two example neurons (bd026b, upper and bd054b, lower) from monkey BD that responded similarly to different views of the same object, and had responses that were different across objects. The images, grouped by object, run along the abscissa; the firing rates (in spikes/s) are shown on the ordinate.

object 4 shown in both colour and greyscale, and to views of the real object. These three sets did not differ significantly from each other. Testing of colour removal from the stimulus was possible for 10 out of the 21 view-invariant neurons reported here. The results are illustrated in Figure 6, which shows the averaged mean firing rate for each cell across the four views of the cell's optimal object, when these were presented in either colour or greyscale. In 7 out of these 10 cases, colour removal had no significant effect on the magnitude of the responses of the neuron to the object to which each neuron responded best. The standard deviation for each cell calculated over the four views is also indicated. It was notable that for the majority of the neurons, the variance calculated across the different views was little affected by the removal of colour, showing that the neurons were producing consistent responses to different views of the

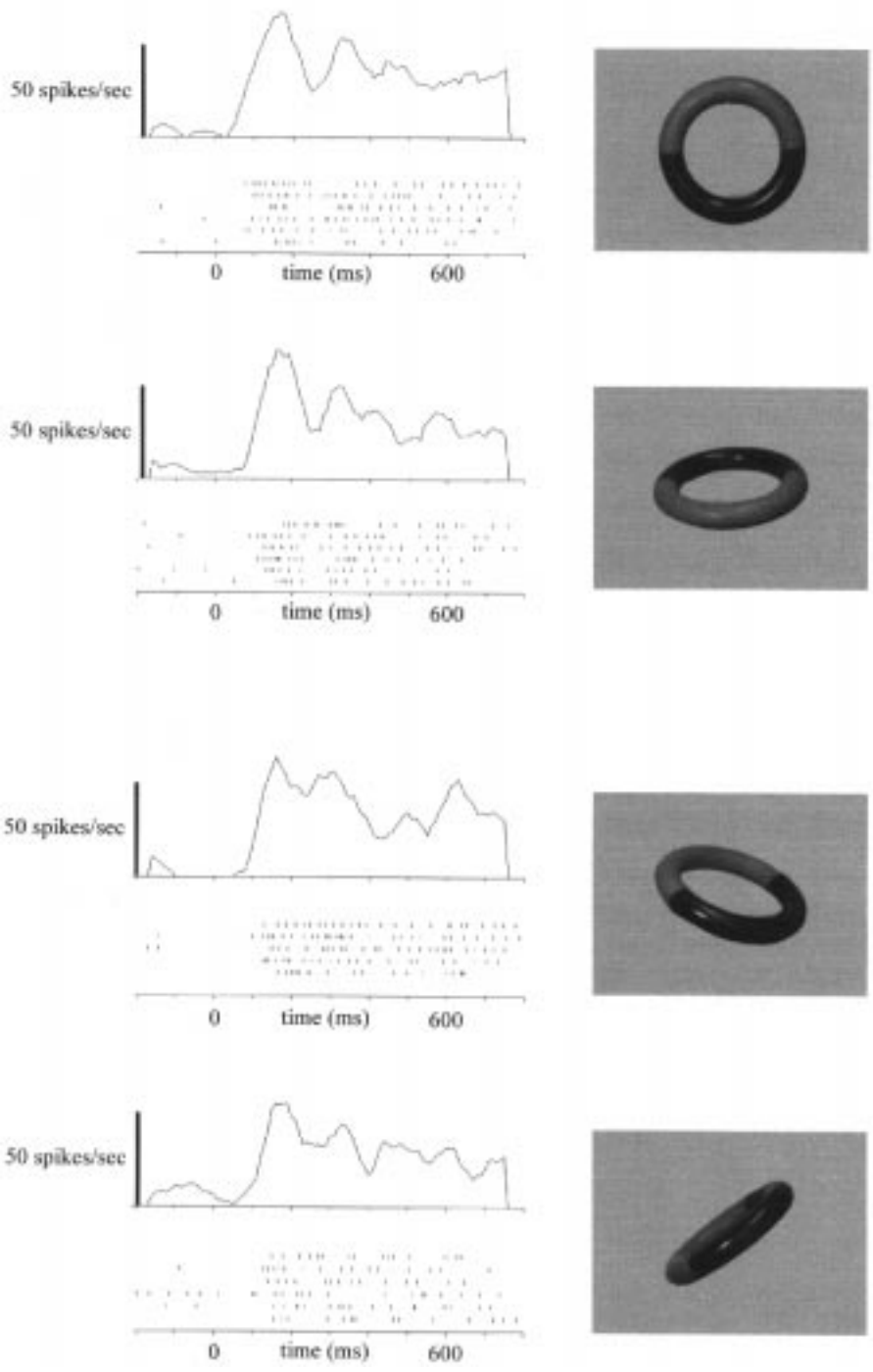

Figure 4. Peristimulus time histograms and associated rasterplots of neuron bd054b's responses to the four views of object 6. The four views are presented adjacent to the appropriate firing rate plots.

same object even when colour was removed. (For only two neurons was the variance significantly affected by removing the colour, Levene's test, $P < 0.05$, and these two cases are indicated on Fig. 6 by a 'v' below the cell number.) The three cases where the magnitude of the response to the greyscale images was significantly smaller ($P < 0.05$) from that to the coloured stimuli are indicated on the graph by an asterisk. In two of these three cases the variance was not affected, indicating that it was not just the colour which enabled the neuron to respond to the different views of an object in a similar way. In another comparison which shows that, in general, colour was not required for the responses of these neurons, we calculated (for the population of 10 cells tested) the correlation between their responses to the coloured and to the greyscale images of the objects. This correlation was 0.816 ($P < 0.004$). The fact that the view-invariant responses of these cells did not depend on colour is further illustrated in Figure 7, which shows peristimulus time histograms made from randomized trials for each of the four different views of object 1

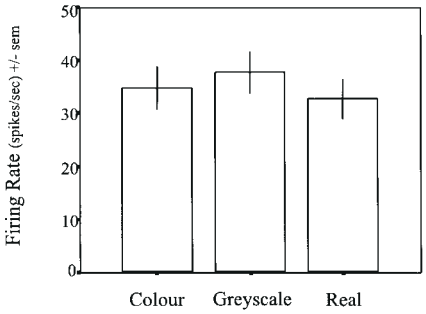

Figure 5. The averaged responses of neuron ay095b to the four views of object 4, shown in both colour and greyscale, and to views of the real object. These three sets of responses were not significantly different.

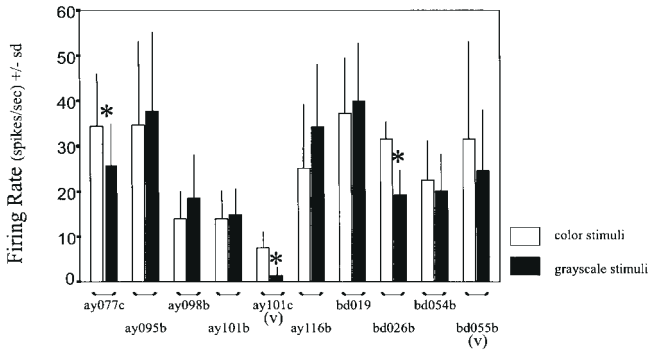

Figure 6. The responses of 10 neurons where the effect of colour removal from the stimuli was possible. The neurons (cases) are plotted on the abscissa (with bars for coloured and greyscale stimuli respectively); the firing rate (in spikes/s) is plotted on the ordinate with the associated standard deviation. In three cases, this manipulation caused a significant decrease in the firing rate of the neuron to the stimuli (*t*-test $P < 0.05$), and these cases are indicated by asterisks '*' on the graph. In addition, the cases in which the variance was significantly altered (Levene's test, $P < 0.05$) are indicated by a 'v' below the neuron number.

when it was presented with colour (above) and in greyscale (below).

For two further neurons, it was possible to test the neuronal response to views of the optimal object that were not included in the regular testing set that the monkey saw on a daily basis. In both these cases, the neurons responded to these unfamiliar screen stimuli at equivalent rates to those produced by the familiar screen stimuli. For example, Figure 8*A* shows the responses of neuron bd054b to three classes of stimuli from object 6: to unfamiliar views (these three views are shown in Fig. 8*B*); to familiar views; and, for comparison, to greyscale versions of familiar views. The neuron's responses across these three classes of stimuli were not significantly different.

In four separate experiments with four different view-invariant neurons, it was possible to investigate whether the view invariance was greater for objects that had been observed repeatedly in the home cage compared to images of objects that had never been seen before as real objects. The hypothesis tested was that the object invariance reflected in the neuronal responses would be greater for the repeatedly seen objects than for the images of different views of objects that had never been seen as real objects. To make this comparison, an object-invariance index was defined for a cell as the variance of the responses within objects divided by the variance of the responses between objects. A low value for this index indicates that a neuron tends to respond similarly to different views of an

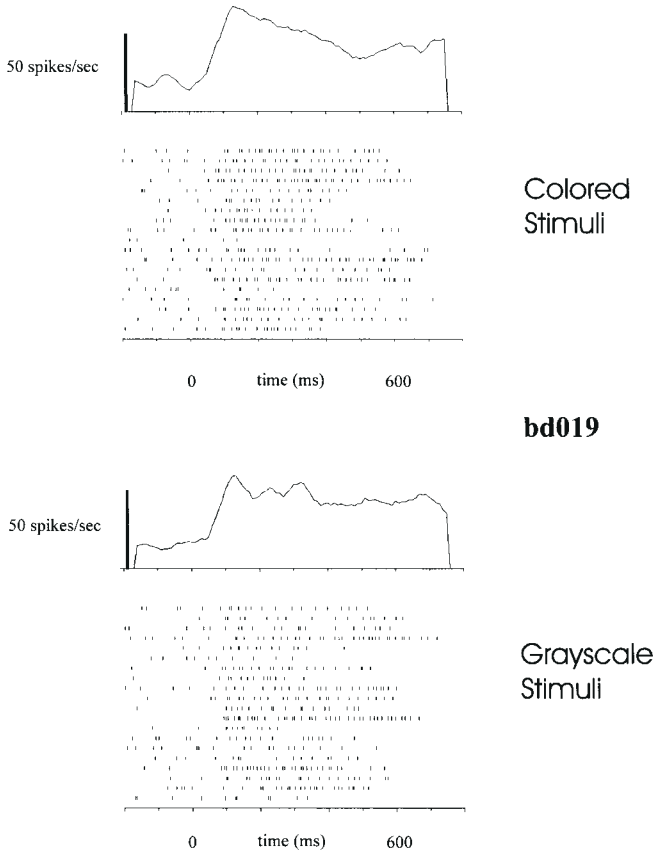

Figure 7. Peristimulus time histograms and associated rasterplots of neuron bd019's responses to coloured and to greyscale views of object 1. These two sets of responses are not significantly different.

object while at the same time having different responses to different objects. The variance of the responses within objects was calculated by taking the variance across the (typically) four views of each object, and averaging this value across objects. The variance of the responses between objects was calculated from the mean firing rates to each object. In all cases the images of the objects were presented in random sequence on the video monitor. The mean-invariance index for these four cells for the cage objects was 0.48 ± 0.22 (SEM), and for the non-cage objects was 2.39 ± 0.95. The invariance index was significantly lower (indicating more invariance) for the cage than for the non-cage objects (Wilcoxon test, $P = 0.034$ one-tailed). Thus the data available from these four experiments suggest that the view-invariant responses of these neurons were due to the fact that the objects had been seen in the home cages from many different views.

Information Theoretic Analyses

The results of the single-cell information analysis are shown in Figure 9. The analysis assessed how much information was available from the firing rate of a neuron in a 500 ms epoch on a single trial about which of the 10 objects was shown. The firing rate data for each object consisted of at least 20 trials of data, i.e. five trials from each of the four views. The average information across all stimuli is shown. The average across the 14 cells in AY was 0.37 bits, and across the 7 cells in BD was 0.53 bits (these means are significantly different [$t(19) = 3.573$, $P < 0.01$] and are illustrated in Figure 10. These average amounts of information

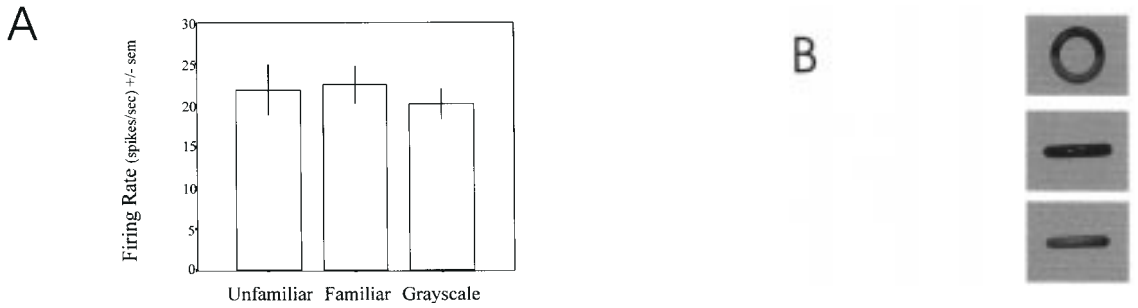

Figure 8. (*A*) The averaged responses of neuron bd054b to three classes of stimuli from object 6: unfamiliar views ($n = 3$); familiar views ($n = 4$); greyscale views ($n = 4$) (see text). (*B*) Greyscale versions of the three unfamiliar views of object 6 used to test neuron bd054b. They were presented in colour in the experiment but are reproduced in greyscale here.

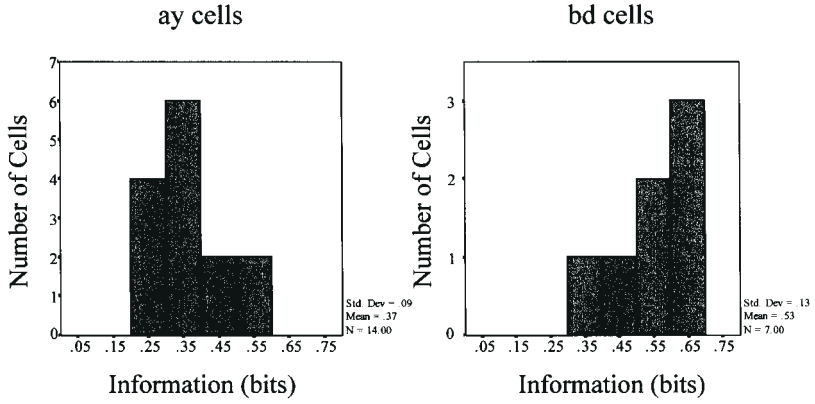

Figure 9. Single-cell information analysis. The average amounts of information (across stimuli) provided by the responses of each cell are indicated in the histograms (shown separately for the two monkeys). Neurons from monkey AY are represented by the histogram on the left, and those recorded from monkey BD are represented in the histogram on the right.

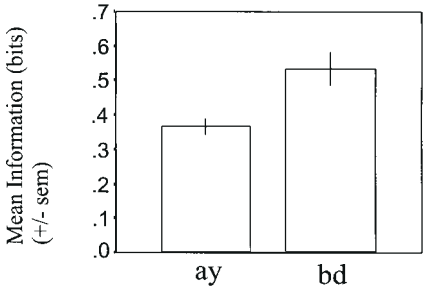

Figure 10. The mean averaged information about object across 14 AY cells and 7 BD cells. These two values are significantly different [$t(19) = 3.573$, P 0.01]. BD cells carry more information about which object is being seen than do neurons from monkey AY.

compare very reasonably with the values of face-selective cells for which of 20 faces was shown (0.36 bits) (Rolls and Tovee, 1995). It should be emphasized here that this information is with respect to object, independently of view. If the neurons had different responses to the different views of an object, then the information about which object had been shown would be much lower. Another way to show how much information the cells provide about objects is to measure the amount of information the cell provides about the most effective object in the set, when it is shown (Rolls *et al.*, 1997a). [The method involves only stopping the procedure required to calculate equation (1) before the summing across stimuli (here objects) is performed.] The mean amounts of information provided about the most effective object in the set were 1.11 bits for AY, and 1.76 bits for BD, and again these means are significantly different [$t(19) = 3.05$, $P < 0.01$]. These values again compare very reasonably with those obtained by Rolls *et al.* (1997a) for the amounts of information provided by face-selective cells about the most effective face in a set of 20 faces (1.8 bits averaged across 14 cells).

The results of the multiple cell information analysis are shown in Figure 11. It is clear that the information about which object was being seen increases as more cells are added to the sample, up to values of ~1.6 bits for AY (14 cells) and ~1.6 bits for BD (7 cells). Moreover, the increase of information is approximately linear over the number of cells in our sample.

Sparseness

Finally, to measure the sparseness of the neural representation of familiar objects we have applied the sparseness measure of Treves and Rolls (1991), calculated according to the following formula for the sparseness a:

$$a = \frac{\left(\sum_{s=1,S} \frac{r_s}{S}\right)^2}{\left(\sum_{s=1,S} \frac{r_s^2}{S}\right)} \qquad (4)$$

where r_s is the firing rate to the sth stimulus in the set of S stimuli. Sparseness ranges from $1/S$, when the cell responds

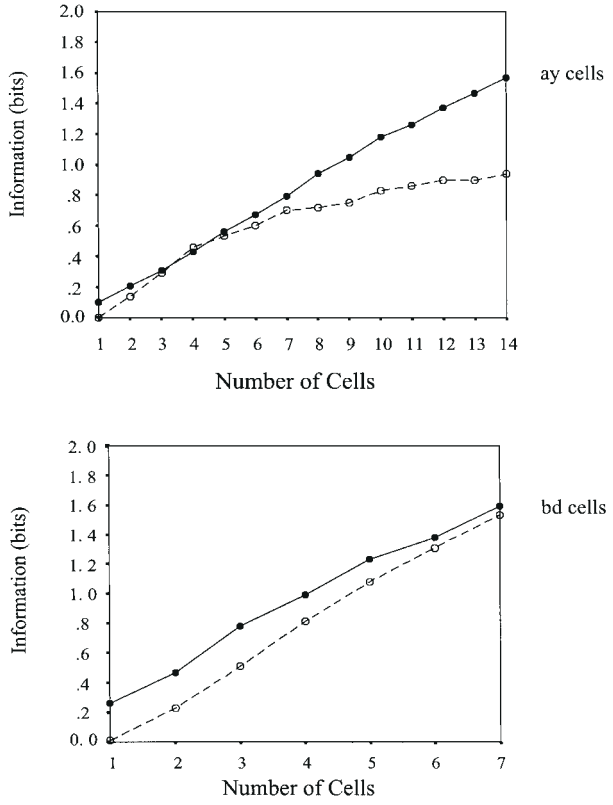

Figure 11. Multiple-cell information analysis. The information provided by different numbers of neurons about which of the 10 objects had been shown. As the number of cells in the sample increases, the information about which object was shown increases in an approximately linear manner. Separate diagrams are provided for the two monkeys, as the object sets were different, and the analysis only makes sense within a single individual. The solid symbols show results with the Bayesian probability estimator (PE) algorithm, and the open symbols with the dot product decoding which is neurophysiologically plausible (see Rolls and Treves, 1998).

Table 1
Response properties of 21 view-invariant neurons

Cell no.	Average information (across all objects)	Optimal object information	Sparseness	Response sparseness	Maximum firing rate (spikes/s)
ay055b	0.379	0.946	0.952	0.789	142
ay055_2	0.291	0.768	0.954	0.594	118
ay057	0.527	1.097	0.721	0.543	50
ay058b	0.514	0.949	0.905	0.65	48
ay074bt	0.458	1.215	0.842	0.605	68
ay077c	0.258	0.76	0.951	0.799	64
ay092b	0.361	1.716	0.503	0.4	26
ay095b	0.366	1.152	0.78	0.477	76
ay098b	0.327	1.027	0.76	0.574	44
ay101b	0.296	1.611	0.874	0.626	30
ay101c	0.351	1.487	0.88	0.604	22
ay109	0.309	0.988	0.95	0.594	96
ay116b	0.27	0.748	0.929	0.722	68
ay118b	0.405	1.088	0.879	0.655	56
AY average	0.365	1.111	0.848	0.616	64.9
bd019	0.541	2.238	0.8	0.653	66
bd021b	0.53	1.221	0.835	0.53	44
bd021c	0.323	0.983	0.891	0.783	40
bd026b	0.693	2.686	0.59	0.554	40
bd043b	0.618	2.044	0.704	0.43	50
bd054b	0.624	2.102	0.423	0.309	44
bd055b	0.404	1.009	0.817	0.648	76
BD average	0.533	1.755	0.722	0.558	51.4
Cell average	0.421	1.325	0.807	0.597	60.4

to only one stimulus, to a maximal value of 1.0, attained when the cell responds with the same rate to all stimuli. The measure works for both neurons with binary responses (e.g. either not firing or firing with a high rate, when it equates to the proportion of active neurons), and with neurons with more continuously distributed firing rate distributions, e.g. exponential-like (Rolls *et al.*, 1997a). The sparseness of the representation of the 10 objects averaged 0.80 across the cells. This compared with a value of 0.65 in face-selective cells responding to a set of 23 faces and 42 non-face stimuli (Rolls and Tovee, 1995). Subtracting the spontaneous firing rate from the firing rate of the neuron, so that the *changes* in firing rate were used in sparseness calculations, gives an average 'response sparseness' of 0.60 (cf. 0.33 of Rolls and Tovee, 1995). The results from the single-cell information analyses, the sparseness calculations and the maximum evoked firing rate for each cell are shown in Table 1.

Comparison with Cells with Non-view-invariant Responses

The full protocol described here was performed for some cells which did not have clear view-invariant responses during the initial presentation of test stimuli, to compare their view-dependent responses, common among cells in the IT, with the relatively view-invariant neurons described in this paper. An example of the responses of one such cell is provided in Figure 12. The neuron responded with very discrepant responses to different views of the same object (see, for example, the responses to objects 7, 9 and 10). Consistent with this, the average information about which object had been seen on any one trial was 0.08 bits. For three cells of this type, the average information about which object had been seen was 0.079 bits. These analyses show that the view-invariant cells described in this paper are indeed making explicit, in their responses, information about which object has been seen invariantly with respect to the view shown.

Recording Sites

Recording tracks in both monkeys were made over an extensive portion of the IT, from the upper and lower banks and fundus of the STS, through the middle temporal gyrus to just lateral to the middle temporal sulcus. The recording sites of the 21 cells reported here are illustrated in Figure 13, and all were recorded in the right hemisphere. As can be seen in the figure, the cells are distributed from lateral to the middle temporal sulcus to the lower bank of the STS, and the investigated area of cortex is indicated by the shaded bounding box on the lower right coronal section.

Discussion

We have shown in this paper that there are cells in IT cortex that respond similarly to different views of the same object, and that these cells convey information about which of 10 objects is being seen, even when there are four views of every object. The experimental design provided evidence that the cells do not simply respond to single features such as colour, for when this was explicitly tested, 7/10 neurons responded similarly to the same views shown in greyscale. Also, as shown in Figures 2 and

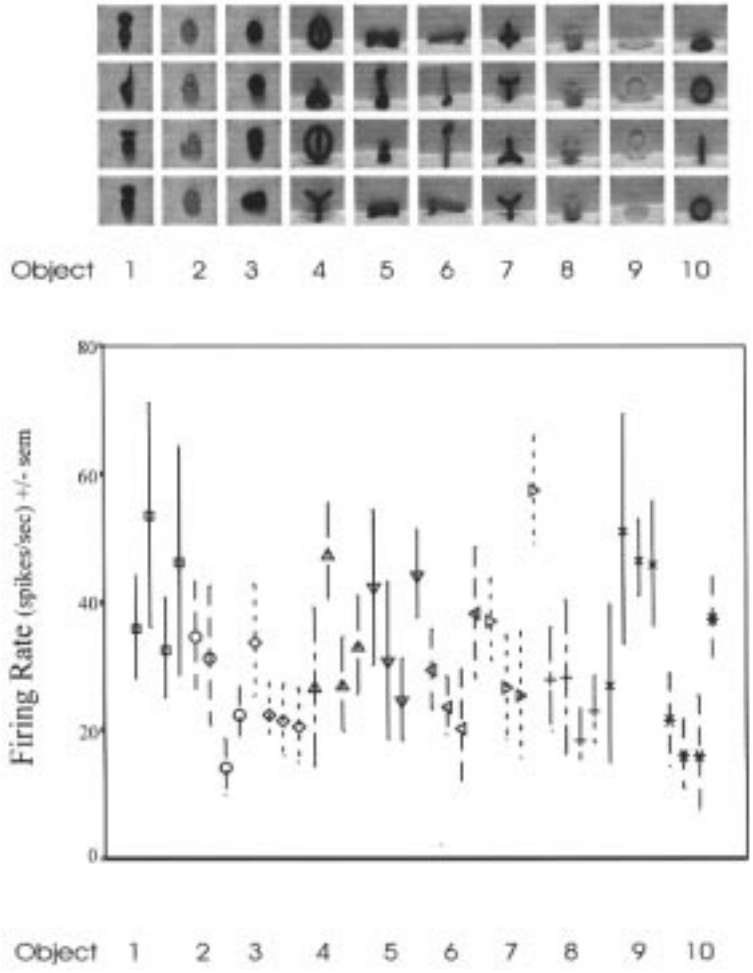

Figure 12. Firing rate profile of neuron ay077b across the 10 objects in set A. The neuron did not have invariant responses to the objects in the set. The images, grouped by object, run along the abscissa; the firing rates (in spikes/s) are shown on the ordinate.

3, the neurons could respond to views of the same object which revealed a different colour. Further, Figures 2 and 3 demonstrate that these neurons could respond to very differently shaped images, with different features, when they were part of one object, and might not respond to superficially similar images and features when they were from another object. This evidence suggests that these cells are not simply feature detectors, responding to a particular feature found in some images, but rather are responding to the object(s) *per se*. It should be noted that the objects used were real-world objects, and so did not offer total control of the features within each image. (We refer to these as real-world objects for, although they might not represent objects that the monkey might encounter in the wild, they are 'real-world' objects of the type seen not only in their home cages but also every day in the laboratory.) In that we used real objects rather than artificially generated images of possible objects, there is a possibility that some features are common to one object and to no other object. However, we took care to choose objects that were rather similar to each other in many respects, including size, surface texture, having few colours and being primarily convex, in order to try to reduce the putative feature space. The firing rate differences between different objects and the firing rate similarities within objects lead us to believe that visual features are not responsible for these invariant responses, and that the neurons are responding when a certain object is being seen. Indeed, we have already noted that 49% of the other visually selective neurons recorded in this investigation did respond differently to the different views of these or other objects, so that feature-based selectivity was a frequent property of many other cells in these brain areas, and this strengthens the suggestion that the object-selective cells described here are different, and are coding for objects as objects, not as a collection of features.

The object representations described here were found with no explicit training. [Training was used by Logothetis *et al.* (1995) and Logothetis and Sheinberg (1996).] The object stimuli were objects that had been present in the home cage, and had been manipulated and viewed from different directions by the monkeys. All of the objects were covered in scratches and bite marks at the end of the experiment, consistent with the view that they had been thoroughly investigated by the monkeys. From the data presented here, the implication is that seeing a real object from different views may enable the visual system to build a view-invariant representation. One hypothesis suggests that the natural transformations of objects would allow different features to become associated together as belonging to the same object. If one feature temporally follows another feature, these features may become associated together (Foldiak, 1991; Rolls, 1992). A neural network model of the visual cortical areas which incorporates a modified Hebbian learning rule that has a

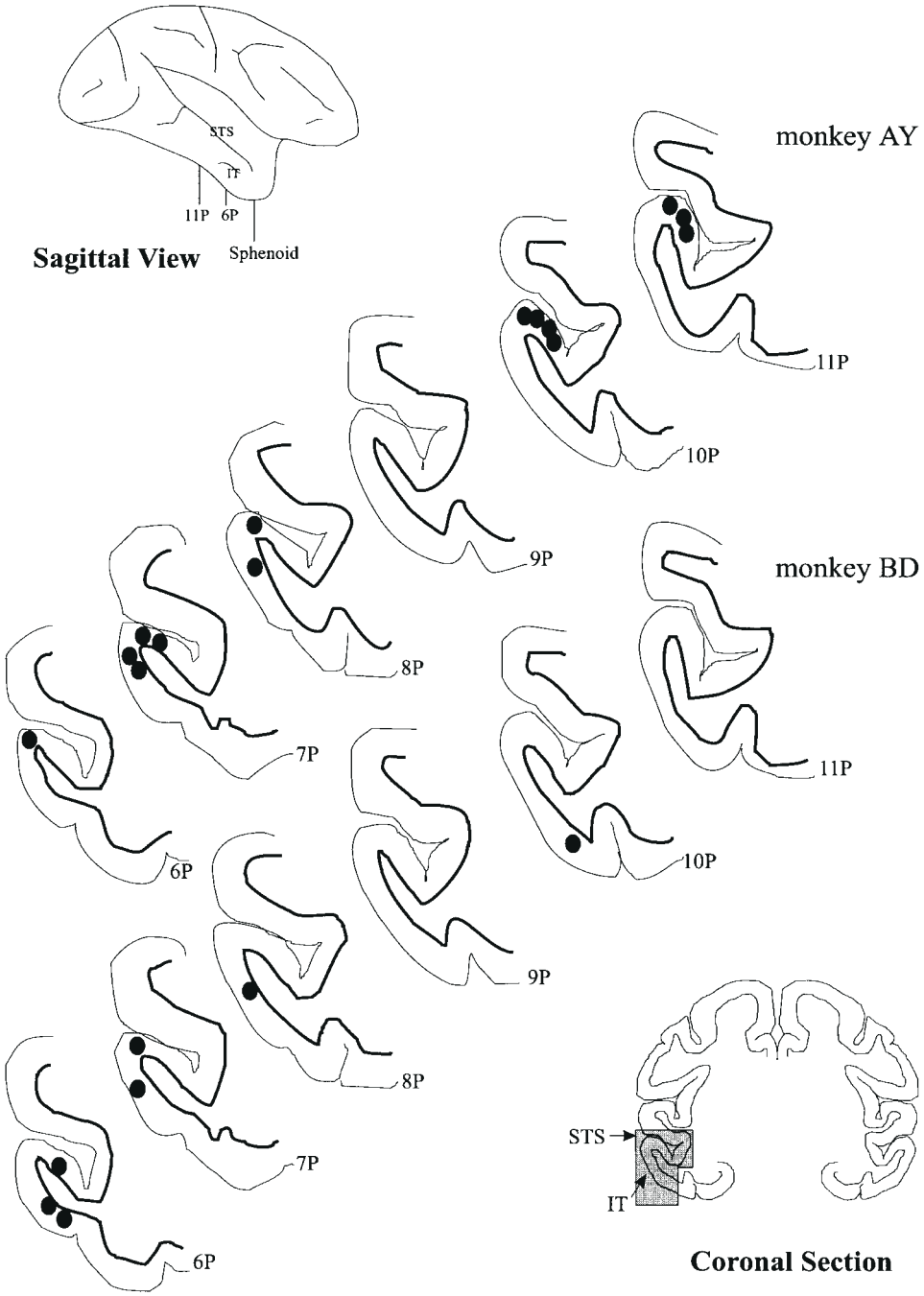

Figure 13. Reconstructed histological sections, for both monkey AY and monkey BD showing (filled circles) the sites at which the view-invariant object neurons were recorded. Numbers alongside the sections indicate the distance (in mm) posterior to the sphenoid bone reference point, and these distances are further illustrated in the upper left of the figure in the sagittal view. A full coronal section is illustrated at the bottom right of the figure, and the area of cortex investigated in this study is indicated by the shaded region encompassing the superior temporal sulcus (STS) and the lateral portion of the inferior temporal gyrus (IT).

short-term memory trace to enable features to be associated together across short times has been built, and can learn view- as well as translation- and size-invariance (Wallis and Rolls, 1997). A biologically plausible computation of this type could be carried out within the temporal cortical processing areas (and in earlier visual cortical areas), perhaps as a product of feedforward hierarchical architecture (Wallis and Rolls, 1997) or by using a trace learning rule in the recurrent collateral connections between neurons (Parga and Rolls, 1998). However, we note that for many objects in the real world it may be possible to respond to an object based on one or several features which do not overlap with features from other objects (see Rolls and Treves, 1998, section 8.8). An example might be a banana and an orange, where the list of features of the banana might include yellow, elongated and smooth surface; and of the orange its orange colour, round shape and dimpled surface. Such objects could be distinguished just on the basis of a list of the properties — no special mechanism is needed for view invariance, because the list of properties is very similar from most viewing angles. However, the objects used in the investigation described here

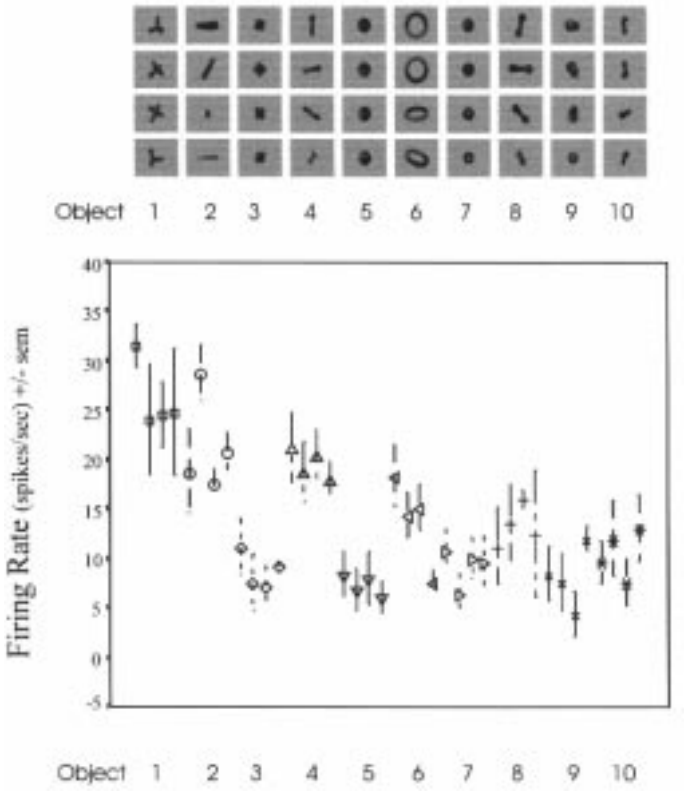

Figure 14. Firing rate profile of neuron bd021b across the 10 objects in set B. There is a degree of clustering at low (suboptimal) rates, especially to objects 3–5 (see text). The images, grouped by object, run along the abscissa; the firing rates (in spikes/s) are shown on the ordinate.

were chosen to minimize object identification by this method, so that the most likely hypothesis is that the monkeys have learned to recognize the cage objects from different views. Consistent with this hypothesis is the fact that in four separate experiments on four different neurons, the view-invariance index indicated more view invariance for views of objects that had been in the home cage than for images of other objects which had never been seen as objects, as described in the Results. Future research in which, for example, recordings are made from neurons while monkeys see unfamiliar objects from different views will be of interest to define the processes and their time course by which the view-invariant object cells described here become established. We also note that if view-invariant neurons are formed from view-based neurons, there might be expected to be many more neurons responding in a view-dependent way (e.g. to features) than in a view-independent way (because each feature or view has to be separately represented). This situation was found in this study, and in an earlier investigation of object-based versus view-based representations of faces (Hasselmo *et al.*, 1989). Further, the anatomical location of the view-invariant cells reported in this paper suggests that these cells do not form a separate anatomical population of neurons in the IT, but rather are intermingled with neurons that are view-dependent and require a certain feature or combination of features for activation. This supports the notion that these view-invariant responses are being formed by associating together the responses of view-dependent visual neurons.

It is also possible that brain areas other than IT and the STS contribute to the formation of view-invariant representations of objects. It is possible, for example, that the associations between different features or views of objects are formed in the perirhinal cortex, which in turn influences the representations in IT by feedback connections. The perirhinal cortex lies in the ventromedial temporal lobe and is reciprocally connected with IT (for review, see Suzuki, 1996) and has been shown to have a functional role that is doubly dissociable from anterodorsal IT cortex (Buckley *et al.*, 1997). Further it has been reported in a tracer study that a small site injection of anterograde tracer PHA-L in the anteroventral part of area TE (TEav) projects diffusely over a wide extent of perirhinal cortex (Saleem and Tanaka, 1996), making the perirhinal cortex anatomically suited for making associations between and across stimuli. Miyashita and colleagues have combined neurophysiological recording in monkey IT with perirhinal cortex lesions, and have suggested that the perirhinal cortex mediates the ability of IT neurons to represent associations between stimulus pairs – these were separated by delays of several seconds in a model of visual–visual association formation in semantic memory (Higuchi and Miyashita, 1996; Miyashita *et al.*, 1996). They taught monkeys pairs of fractal patterns and then recorded from neurons in IT. After training it was found that the responses of these neurons to the fractal stimuli tended to be correlated, i.e. a neuron firing to one of the patterns would have a significant chance of also responding to the other member of the stimulus pair. The perirhinal cortex was then unilaterally ablated and the monkeys were taught a further set of fractal pairs. When recording from IT in the lesioned hemisphere, they found that the neuronal response correlation had been abolished for both the pre-operative and post-operative fractal pairs. However, these studies were carried out using extensive training to associate the images with each other, whereas there was no such training in the current study. Ablation studies have implicated the perirhinal cortex in object knowledge and object memory (Buckley and Gaffan, 1997; Buckley *et al.*, 1997), although once again extensive training procedures were used to pair object stimuli together and with rewards. An important focus of future experiments is to measure the effects of perirhinal cortex lesions on the formation of view invariant representations of new objects.

Our results do not support a local or 'grandmother cell' coding scheme as suggested by Barlow (1972, 1995). Such coding would require that the view-invariant cells respond to all views of one object and only to that object. This is clearly not the case, and is quantified by the sparseness measure which indicates a distributed representation of the object information. The neurons described here may respond to more than one of the objects in the set, and, in addition, may respond to some views of other objects. Further, the information theoretic analysis shows that as the number of cells increases, the information about what object is being seen increases approximately linearly – and so the number of stimuli (objects) that can be discriminated increases exponentially. A grandmother cell coding scheme would lead to only a linear increase in the number of discriminable stimuli (Rolls and Treves, 1998). Thus the present results indicate that the coding of objects, as well as that of faces, in the IT uses a sparse distributed representation, with associated advantages of generalization, graceful degradation and an ability to represent a great deal of information in any random subset of the cells. Our findings also show that a great deal of information is provided by the firing rates of a population of cells, without the need to take account of the relative time of firing of cells (cf. Singer and Gray, 1995).

Another result of this experiment is that BD neurons carry

more information (both in terms of average information about object and in terms of information about the optimal object in the set) about the respective object set than do AY neurons. We believe that this difference reflects the fact that the object set given to monkey BD was richer in terms of visual features than that given to monkey AY, and thus there is more information to be coded by the neurons. The other response characteristics of the neurons do not differ across the two populations — the mean sparseness, mean response sparseness and the mean maximum firing rate are not different across the data for the two monkeys.

The experimental method used enabled us to analyse neuronal responses to many different stimuli, as opposed to finding an optimal stimulus/feature/colour/shape for the cell and then altering this (Tanaka, 1993). One interesting result of this is shown in Figure 14. Neuron bd021b's optimal stimulus is view 1 of object 1, but it is apparent that there is 'clustering' of lower level firing rates to objects 3–5. These responses are greater than the spontaneous firing rate of the cell. Neurons coding for objects in this way might use their entire dynamic range for signalling, and be able to transmit information much more efficiently. Rolls *et al.* (1997a) have shown that much information is transmitted in the low (sub-optimal) firing rates of IT neurons, and recent work has been undertaken to investigate the efficiency of the coding schemes used by IT neurons (Booth *et al.*, 1997).

An interesting issue for future research is exactly how these view-independent neural responses are formed, and indeed what their contributions to perception might be. One hypothesis is that they contribute to the rapid recognition of familiar objects by simplifying the output of the visual system. However, further work investigating interactions within and between the inferotemporal and perirhinal cortices is required before we can answer these questions.

Notes

This work was supported by a Medical Research Council (UK) grant PG8513790 to ETR and a Wellcome Trust studentship to M.C.A.B.

Address correspondence to Professor E. T. Rolls, Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK. Email: edmund.rolls@psy.ox.ac.uk.

References

Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology? Perception 1:371–394.

Barlow HB (1995) The neuron doctrine in perception. In: The cognitive neurosciences (Gazzaniga MS, ed.), pp. 415–435. Cambridge, MA: MIT Press.

Booth MCA, Panzeri S, Treves A, Tovee MJ, Rolls ET (1997) Efficiency of encoding of natural scenes by neurons in the inferior temporal cortex. Soc Neurosci Abstr 23:2063.

Bruce C, 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.

Buckley MJ, Gaffan D (1997) Impairment of visual object-discrimination learning after perirhinal cortex ablation. Behav Neurosci 111:467–475.

Buckley MJ, Gaffan D, Murray EA (1997) Functional double dissociation between two inferior temporal cortical areas: perirhinal cortex versus middle temporal gyrus. J Neurophysiol 77:587–598.

Cover TM, Thomas JA (1991) Elements of information theory. New York: Wiley.

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051–2062.

Farah, MJ (1990) Visual agnosia. Cambridge, MA: MIT Press.

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

Foldiak P (1991) Learning invariance from transformation sequences. Neural Comput 3:194–200.

Fujita I, Tanaka K, Ito M, Cheng K (1992) Columns for visual features of objects in monkey inferotemporal cortex. Nature 360:343–346.

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

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

Higuchi S, Miyashita Y (1996) Formation of mnemonic neuronal responses to visual paired associates in inferotemporal cortex is impaired by perirhinal and entorhinal lesions. Proc Natl Acad Sci USA 93:739–743.

Humphreys GW, Riddoch MJ (1987) To see but not to see: a case study of visual agnosia. London: Erlbaum.

Ito M, Tamura H, Fujita I, Tanaka K (1995) Size and position invariance of neuronal responses in monkey inferotemporal cortex. J Neurophysiol 73:218–226.

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

Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J Neurophysiol 71:856–867.

Logothetis NK, Sheinberg, DL (1996) Visual object recognition. Annu Rev Neurosci 19:577–621.

Logothetis NK, Pauls J, Poggio T (1995) Shape representation in the inferior temporal cortex of monkeys. Curr Biol 5:552–563.

Mikami A, Nakamura K, Kubota K (1994) Neuronal responses to photographs in the superior temporal sulcus of the rhesus monkey. Behav Brain Res 60:1–13.

Miyashita Y, Okuno H, Tokuyama W, Ihara T, Nakajima K (1996) Feedback signal from medial temporal lobe mediates visual associative mnemonic codes of inferotemporal neurons. Cog Brain Res 5:81–86.

Nakamura K, Matsumoto K, Mikami A, Kubota K (1994) Visual response properties of single neurons in the temporal pole of behaving monkeys. J Neurophysiol 71:1206–1221.

Panzeri S, Treves A (1996) Analytical estimates of limited sampling biases in different information measures. Network 7:87–107.

Parga N, Rolls ET (1998) Transform invariant recognition by association in a recurrent network. Neural Comput (in press).

Perrett DI, Oram MW, Harries MH, Bevan R, Hietanen JK, Benson PJ, Thomas S (1991) Viewer-centred and object-centred coding of heads in the macaque temporal cortex. Exp Brain Res 86:159–173.

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

Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves, MA (1985) Visual cells in the temporal cortex sensitive to face view and gaze direction. Proc R Soc Lond B Biol Sci 223:293–317.

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

Rolls ET (1994) Brain mechanisms for invariant visual recognition and learning. Behav Processes 33:113–138.

Rolls ET (1995) Learning mechanisms in the temporal lobe visual cortex. Behav Brain Res 66:177–185.

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, Tovee MJ (1995) Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. J Neurophysiol 73:713–26.

Rolls ET, Treves A (1998) Neural networks and brain function. Oxford: Oxford University Press.

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

Rolls ET, Treves A, Robertson RG, Georges-Francois P, Panzeri S (1998) Information about spatial view in an ensemble of primate hippocampal cells. J Neurophysiol 79:1797–1813.

Rolls ET, Treves A, Tovee MJ (1997a) The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex. Exp Brain Res 114:149–162.

Rolls ET, Treves A, Tovee MJ, Panzeri S (1997b) Information in the neuronal representation of individual stimuli in the primate temporal visual cortex. J Comput Neurosci 4:309–333.

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

Saleem KS, Tanaka K (1996) Divergent projections from the anterior inferotemporal area TE to the perirhinal and entorhinal cortices in the macaque monkey. J Neurosci 16:4757–4775.

Shannon CE (1948) A mathematical theory of communication. AT&T Bell Lab Tech J 27:379–423.

Singer W, Gray, CM (1995) Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18:555–586.

Suzuki WA (1996) The anatomy, physiology and functions of the perirhinal cortex. Curr Opin Neurobiol 6:179–186.

Tanaka K (1993) Neuronal mechanisms of object recognition. Science 262:685–688.

Tanaka K (1996) Inferotemporal cortex and object vision. Annu Rev Neurosci 19:109–139.

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

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

Treves A, Panzeri S (1995) The upward bias in measures of information derived from limited data samples. Neural Comput 7:399–407.

Treves A, Rolls ET (1991) What determines the capacity of autoassociative memories in the brain? Network 2:371–397.

Wachsmuth E, Oram MW, Perrett DI (1994) Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. Cereb Cortex 4:509–22.

Wallis G, Rolls ET (1997) Invariant face and object recognition in the visual system. Progr Neurobiol 51:167–194.

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