Responses of Macaque Inferior Temporal Neurons to Overlapping Shapes

Marcus Missal, Rufin Vogels and Guy A. Orban

Laboratorium voor Neuro-en Psychofysiologie, Faculteit Geneeskunde, KULeuven, Leuven, Belgium

We tested whether macaque inferior temporal neurons can signal the presence of their preferred shape independently of other shapes simultaneously present, by comparing the responses and selectivity of TE neurons to shapes (figures), either presented in isolation or overlapping another shape (background). The two overlapping shapes differed in color or texture and thus were easily segmentable. We found that the response and selectivity of TE neurons to the figure can be dramatically altered, most commonly reduced, when the figure is overlapping the background. This reduction in response was also present when the monkey was required to actively discriminate the figures from varying backgrounds during recording. The level of suppression depended on the shape of the background and on whether or not figure and background can be discriminated. These results indicate that responses of TE cells are not only determined by the properties of the figures but also are influenced by the properties of stimuli in the background.

Neurons in macaque inferior temporal cortex are selective for shape (Gross *et al.*, 1972), color (Komatsu *et al.*, 1992) and texture (Sary *et al.*, 1995), and display a considerable amount of invariance to changes in the retinal image when varying position (Schwartz *et al.*, 1983), size (Schwartz *et al.*, 1983; Rolls and Baylis, 1986), defining visual cue (Sary *et al.*, 1993) or the degree of partial occlusion (Kovacs *et al.*, 1995) of the stimulus (for review see Vogels and Orban, 1996). The selectivity for object features and the well-documented impairments in object recognition after inferior temporal (IT) cortex lesions (Dean, 1976) indicate that this part of the brain plays a substantial role in object recognition (for review see Logothetis and Sheinberg, 1996).

Shapes are an important and defining property of an object; for example, different objects of the same color and texture can be readily discriminated by their shape. Thus, shapes have been widely used to study mechanisms of object coding and IT neurons have been found to respond selectively to shapes that constitute objects; i.e. a shape can be the 'critical feature' of an IT neuron (for review see Tanaka, 1996). Since IT neurons can signal the presence of their preferred shape or 'critical feature' within their receptive field, an object can be coded by the combined activity of units, each of which responds to a different shape (part) belonging to the object. Such a view of how objects are coded assumes, at least in its computationally simplest version, that the selectivity of the neuron is independent of the presence of other shapes, or features, in the receptive field of the neuron. Indeed, only part of a natural object may contain the 'critical feature' and thus will be surrounded by the other shapes within the same object (e.g. the nose shape on a tiger's face). Furthermore, in a natural visual environment, features of other objects will neighbor the shape that the cell is selective for (e.g. in cases of one object partially occluding another). If IT neurons merely signal the presence of their preferred shape, then one would expect that their response would be invariant for the presence of such neighboring shapes which may belong to either the same or a different object.

Response invariance for the presence of a second, neighboring shape was tested in the present series of experiments by comparing the responses and selectivity of IT neurons to shapes presented either in isolation or overlapping another shape. The two overlapping shapes differed in color or texture and thus were easily segmentable. In fact, a response invariance for the presence of neighboring shapes implies that the neurons can also segment the two shapes. Thus, the same experiments also addressed the complementary question of whether IT neurons can segment two overlapping shapes which can be segregated perceptually by virtue of other cues such as color or texture. The effect of overlapping shapes was studied in two behavioral paradigms: a passive fixation and an active shape discrimination task. In the latter condition, active segmentation of the two shapes was required.

Some of the results have appeared previously in abstract form (Missal *et al.*, 1995).

Material and Methods

Subjects and Surgery

Two juvenile male macaque monkeys (*Macaca mulatta*) were used as subjects (T and J). For the experimental sessions, the monkeys were deprived of water but received dry food *ad libitum*, supplemented by fruits as necessary. A scleral search coil was implanted on one eye (Judge *et al.*, 1980) and a stainless steel peg was cemented to the skull for head fixation. Single-cell recordings were made through a recording chamber placed dorso-laterally to the IT cortex. Positioning of the recording chamber was guided by magnetic resonance anatomical images that were taken before surgery. Accordingly, the center of the recording chamber was placed over the left hemisphere 1.7 cm anterior to the auditory meatus and 2.4 cm lateral in monkey T, and 1.8 cm anterior and 2.4 cm lateral in monkey J. The chamber was tilted by 10° to allow oblique penetrations into the IT cortex. All surgical procedures were performed under deep anesthesia and aseptic conditions.

At the end of the recording sessions, several penetrations were made in the cortex of monkey T using electrodes covered with a lipophilic fluorescent dye, carbocyanine (1,1′-dioctadecyl-3,3,3′,3′-tetramethyl-indocarbocyanine perchlorate; DiI) (Honig and Hume, 1989) following the procedure described by Snodderly and Gur (1995). The monkey was subsequently killed with an overdose of Nembutal and perfused with fixative. Recording sites were reconstructed by identifying the tracks made by the DiI-stained penetrations on coronal sections of the brain. Also, after some of the recording sessions, radiographic images of the skull with the electrode *in situ* were obtained for both monkeys, so that the positioning of the chamber and electrode could be compared between the two monkeys.

The procedures conformed to the guidelines established by the NIH for the care and use of laboratory animals. Adequate measures were taken in order to minimize pain or discomfort.

Cerebral Cortex Dec 1997;7:758–767; 1047–3211/97/$4.00

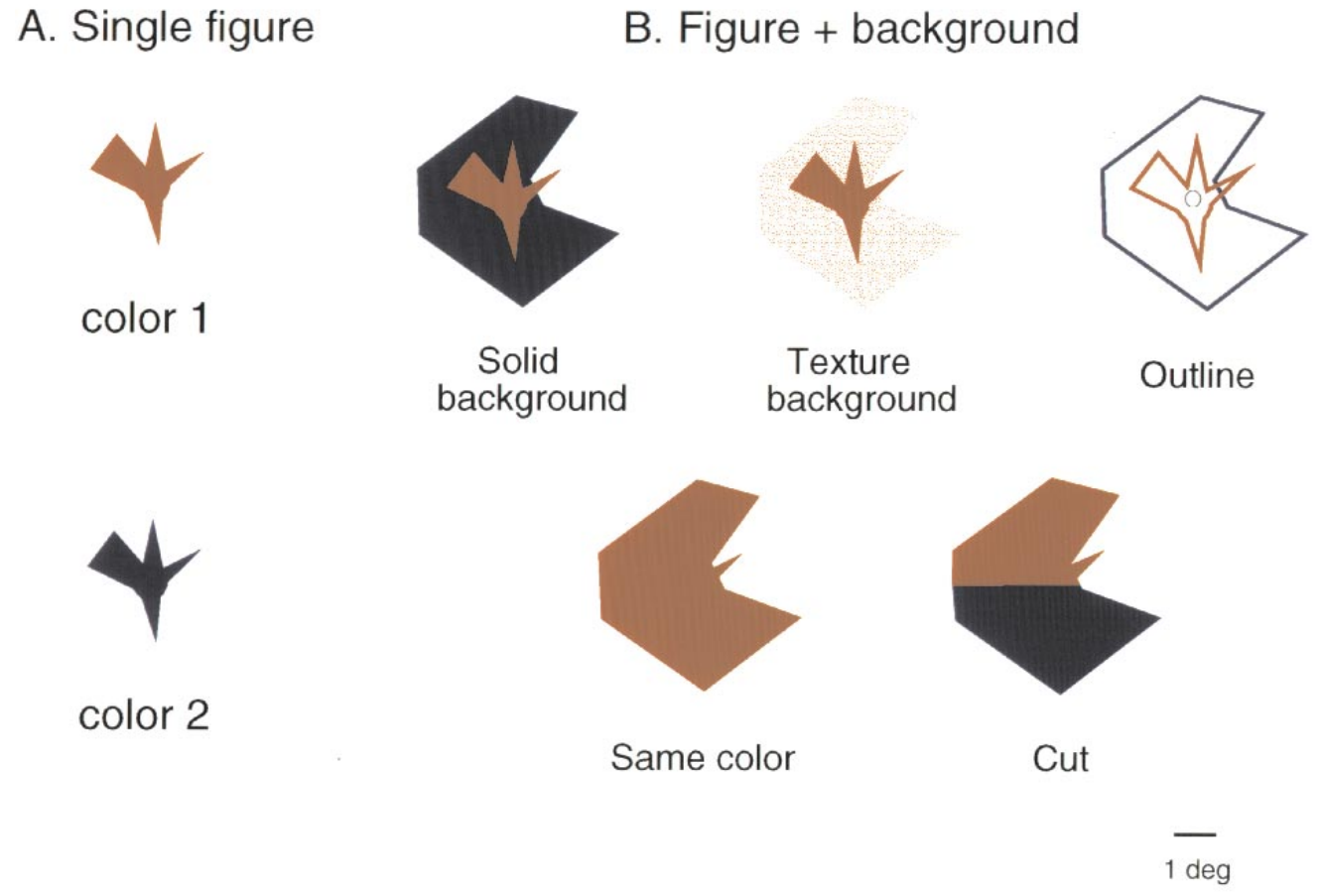

Figure 1. Drawings of the different types of stimuli. (*A*) Single figures were drawn in two different colors: color 1 was the best color for evoking a response from the cell; color 2 was another color. Colors tested were white, red, green and blue. (*B*) Figure + background. Solid background: the figure in color 1 was overlapping the background shape in color 2. Figure and background were solid. Texture background: the figure was solid (color 1) and the background was filled with a random dot texture pattern (50% density) of the same color (color 1). Outline: the figure in color 1 and the background in color 2 were drawn with an outline. Same color: the figure and the background had the same color. Cut: the figure was superimposed on the background but the combination figure + background was cut by an imaginary horizontal line and the shapes in the upper part were put in color 1, in the lower part in color 2. All stimuli were displayed on top of the fixation point which remained visible only with outline stimuli. Shapes were centered on the fixation position, indicated by a small circle in the display of the outline stimuli.

Apparatus

The monkeys were seated in a primate chair, with the head fixed and facing a computer screen. Eye position was sampled at 200 Hz and displayed on a monitor by a PC. Single cell recordings were made with parylene- and kapton-coated tungsten electrodes (World Precision Instruments, impedance 1.5–2.0 MΩ). Electrodes were lowered into the brain using a Narishige microdrive. The signal was amplified and filtered and single cells isolated using a time-amplitude window discriminator (Bak Electronics Inc., model DIS-I). The times of spike occurrence, stimulus and behavioral events were stored and displayed as raster and peristimulus time histograms by a second PC. Stimuli were generated by a Silicon Graphics Indy computer and displayed on a Sony Trinitron screen (1024 × 1024 pixels; width: 39°, height: 29.5°) 58 cm from the monkey's eyes.

Stimuli

In each recording session, we used a set of 30 different shapes that were extracted from a computer-generated set of 4096 polygonal shapes (Fig. 1). The median of the largest diameter (distance between the two extreme apices) of the shapes of the recording set was 3.0° (quartiles, 2.7–3.3). All the shapes except two (circle and square) were asymmetrical.

Shapes were displayed on a uniform black screen (luminance 0.02 cd/m^2) either as a uniform surface of a certain color and luminance (solid filled shapes) or as an outline (see Fig. 1). Luminance of the solid, filled shapes depended on the color: 71.9 cd/m^2 for white, 18.8 cd/m^2 for red, 49.7 cd/m^2 for green and 5.4 cd/m^2 for blue. We did not attempt to match the luminance of these different color stimuli since color was used as a segmenting cue in the present study, and it was not the purpose to measure the selectivity for color *per se*. For outline shapes, the luminance was the same inside and outside the shape outline (black screen). The thickness of the outline was 0.2° and the luminance of the outline was the same as the luminance of the colors used for solid shapes. Shapes could also be filled with a random-dot texture pattern (density 50%). Average luminance was then 2.4 times lower than that for the solid shapes.

Shapes were presented either alone or in combination with a second shape (see Fig.1). In the latter case, one shape, the figure, overlapped a larger shape, the background shape. The size of the background shape was ~1.6 times the size of the figure (see Fig. 1). The two shapes differed in color or texture, and could thus be easily segmented. Centering of the shapes was achieved by aligning their center of mass on the fixation point.

Stimulus Sequences and Behavioral Paradigms during Recordings

Fixation Paradigm

During the fixation task, the monkeys had to foveate a white fixation target (size: 22 min arc) for at least 700 ms. Gaze had to be within a square fixation window of 1.5 × 1.5°. The stimulus was then displayed for 300 ms, throughout which the monkey had to maintain fixation. The stimulus

was a shape presented either with or without a background shape (see below). The stimulus was replaced by the fixation target after a variable delay. The reward was given after the disappearance of the stimulus.

Each cell isolated was first tested with the recording set of 30 shapes. Shapes were presented in isolation in each of the four colors (white, red, green, blue). If a cell responded to at least one of the shapes, it was tested further. From the set of 30 shapes, two shapes driving the cell and two shapes which were less effective were selected. These four shapes, which differed from neuron to neuron, will be called *figures*. Also, two shapes not driving the cell when presented in isolation were selected as background shapes and will be called *backgrounds*. Presentations of the four figures either in isolation in two different colors or combined with one of the two backgrounds were randomly interleaved in a test with 16 conditions. Four conditions consisted of the presentation of the four single figures in the optimum color (color 1). Four other conditions consisted of the presentation of the same figures in another color (color 2) chosen amongst the three other colors tested. The next four conditions consisted of the four figures in color 1 overlapping one of the background shapes in color 2. The last four conditions consisted of the presentation of the four figures in color 1 overlapping the other background shape in color 2.

Discrimination Paradigm

After 700 ms of fixation the stimulus was presented as in the fixation paradigm. However, after 300 ms of stimulus presentation, two target spots were shown, on either side of the figure, at a distance of 8°. For five figures the monkeys had to make a saccadic eye movement to the left spot and for the other five a saccade to the right spot in order to obtain a juice reward. These 10 shapes were randomly selected from the set of 30 shapes used in the fixation task. Stimulus, target spots and fixation spot were turned off when the monkey made the saccade. Trials in which either a saccade occurred before target spot onset or outside one of the two electronic target windows were not rewarded (abort trials).

The same two subjects were trained in this task as follows. Once they learned to discriminate single figures, backgrounds were added to the figures. These backgrounds were unlearned shapes randomly selected from the initial set of 4086 shapes and were randomly changed from trial to trial in single training sessions. The latter prevented the subjects from basing their discrimination on particular figure–background combinations instead of on the figures per se. Both animals rapidly learned to segment the figure from the background and their discrimination performance was ~90% correct with the randomly changing backgrounds. The performance during the recordings was a 97.9 ± 0.4% correct response for single figures. For the two figure–background combinations the performances were similar, 89.6 ± 1.4% and 87.9 ± 1.4%, indicating that the animals were segmenting the figure from the background during the recordings.

For each neuron, we first recorded the response to the 10 learned shapes presented in isolation. From the set of learned shapes, two shapes driving the cell well and two shapes which were less effective were selected (figures), in order to test shape selectivity. Each of the four figures were presented either alone in two different colors or combined with two background shapes in a test with 16 interleaved conditions (see Fixation Paradigm). The background shapes were selected from the set of 4086 shapes (4096 – 10 learned shapes). In order to avoid a response bias, two of the four figures were associated with a leftward saccade. The neuronal response of the cell was recorded for both correct and incorrect discriminations. Whatever the task, at least 10 trials of each of the conditions were carried out.

Data Analysis

Off-line spike counts were computed trialwise with a 300 ms bin that started 50 ms after shape onset. Net responses were calculated by trialwise subtraction of the activity during a fixation period of the same duration as the stimulus time window, but just preceding shape onset. Cells were included in the analysis if there was a statistically significant shape-selective response to the four single figures tested in color 1. This was tested by an analysis of variance (ANOVA; Kirk, 1968). Tests were considered as significant if the corresponding type I error was smaller than 0.05. A cell was defined as color selective if the net response of the cell to the preferred figure in the best color (color 1) was at least twice the net response to the same figure in the other color tested (color 2).

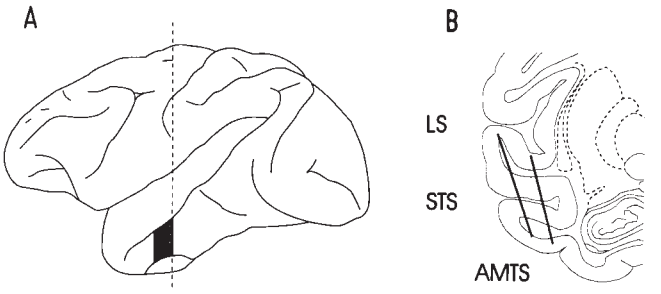

Figure 2. Recordings sites in monkey T. (*A*) Lateral view of the left hemisphere showing the range of recording positions (black). (*B*) Coronal section made at the level indicated by the vertical dashed line on the lateral view of the brain in *A*. The track of two penetrations are indicated by the lines. STS: superior temporal sulcus; AMTS: anterior middle temporal sulcus; LS: lateral sulcus.

A response index (RI) was used to quantify the effect of a background on the shape-selective response of a neuron in order to compare different cells or the same cell tested with different stimuli (texture, size, etc.). The RI was obtained by subtracting the average net response to the combination of preferred figure and background from the average net response to the preferred figure alone, and dividing this difference by the sum of the two responses. When more than one background was used, the combination figure–background with the strongest reduction in response was used to compute the response index. A value of + 0.33 corresponds to a 50% reduction in response.

Results

We recorded from 139 shape selective cells in the IT cortex of two monkeys (91 cells in monkey T and 48 cells in monkey J). Histological analysis of the recording sites in monkey T showed that the recordings were restricted to area TE and excluded area TEO (see Fig. 2). Neurons in both the lower bank of the superior temporal sulcus and the lateral convexity of the inferior temporal gyrus were recorded. Comparing radiographic images of electrode penetrations in the two monkeys suggested that neurons were also located in area TE in the other monkey and did not include the more posterior visual area TEO.

Effect of Background on the Neuronal Response and Shape Selectivity

The effect of adding a background shape to the figure is shown in Figure 3 for a single shape-selective neuron. This neuron responded similarly to both colors, the response to the preferred figure in white (first column, second row in Fig. 3) being 77% of the response to the preferred figure in red (first column, first row in Fig. 3). The last two rows at the bottom of Figure 3 show the response when a background was added to the figures, indicating a strong reduction of the response when the background shape was added. On average, the response of the cell was reduced by 60% for the preferred figure. Also, shape selectivity was strongly affected: the cell did not respond to the second figure when a background was added and the response to the fourth figure was increased with the first background and reduced with the other (Fig. 3). Thus, the selectivity of the cell changed when a background shape was present and this modulation of shape selectivity depended on the background shape.

Such a dramatic effect by a background shape was a typical finding, as shown in Figure 4, which plots the average normalized response to single figures and figures combined with each of two different backgrounds. It summarizes the effects of

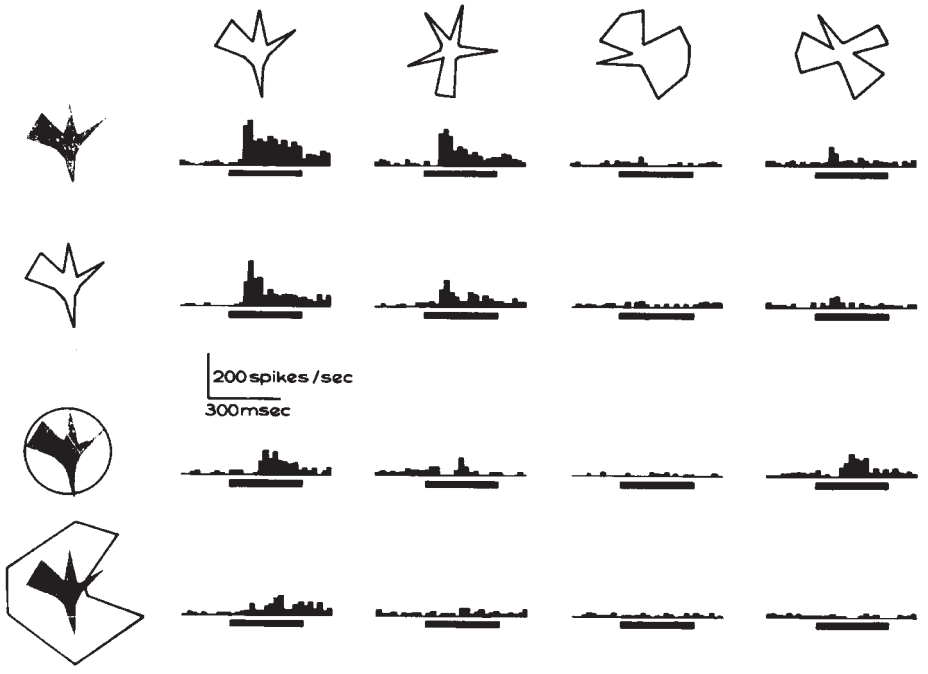

Figure 3. Selectivity and effect of background shape on the response of a shape selective TE unit (monkey T). Peristimulus time histograms for each of the four single figures in one of two colors (upper two rows) and figures with two different background shapes (lower two rows). The figures used to study this cell are drawn above the first row of histograms. The different stimulus configurations are represented, with the first figure given as an example, on the left part of the Figure. Single figures were red (first row) or white (second row). The red color is represented by black on the Figure. Figures were overlapping two different background shapes (background 1, third row; background 2, fourth row). Time of stimulus presentation (10 trials per condition) is indicated by the horizontal bars. Calibration bars are indicated.

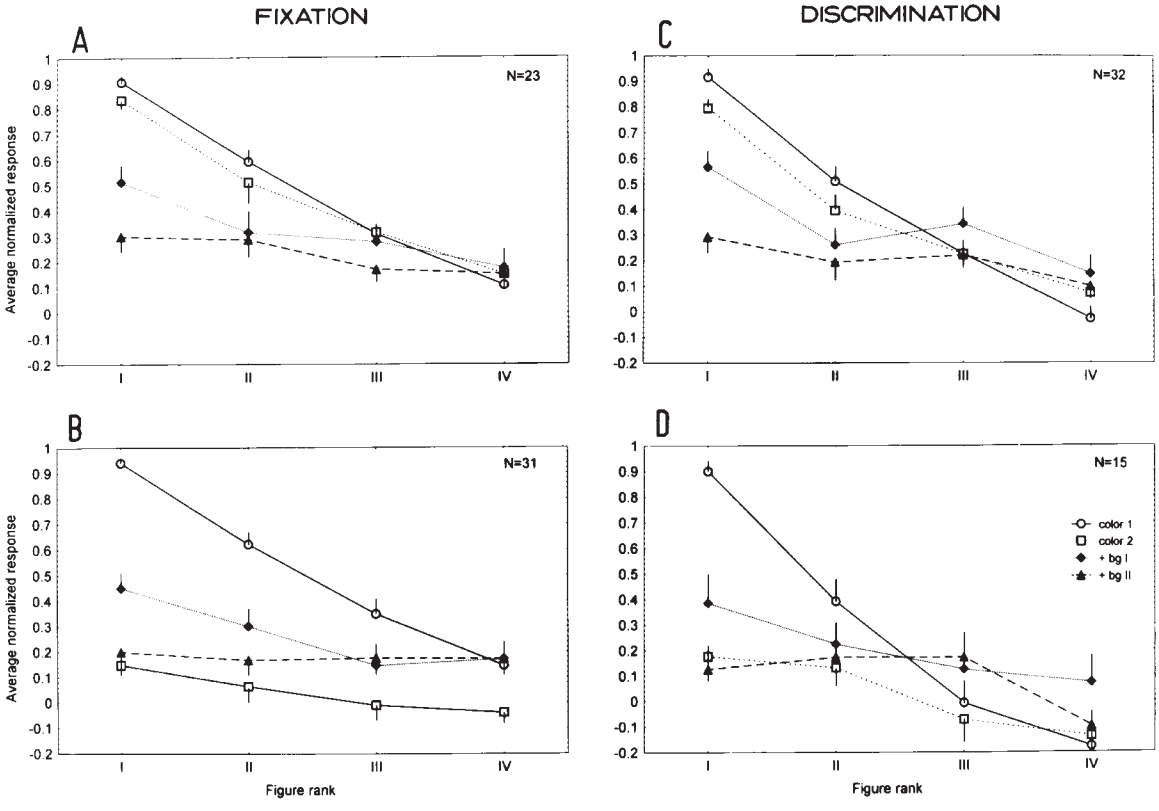

Figure 4. Effect of adding a background shape on the selectivity of TE units. (*A*) Fixation task, shape selective cells. Average normalized response as a function of figure rank for single figures and figures + backgrounds; 23 neurons tested. (*B*) Fixation task, shape and color selective cells; 31 neurons tested. (*C*) Discrimination task, shape selective cells; 32 neurons tested. (*D*) Discrimination task, shape and color selective cells; 15 neurons tested. The figure ranks were determined using the response to single figures. Normalization was done on each unit separately with respect to the strongest response of the cell. Inset indicated the four different conditions.

background shapes for neurons recorded in the fixation (Fig. 4*A*, *B*) and discrimination (Fig. 4*C*, *D*) paradigms. For each paradigm, the data are pooled separately for shape-selective (Fig. 4*A*, *C*) and shape- and color-selective (Fig. 4*B*, *D*) cells. For each neuron we ranked the four single figures according to their net response in the best color; for example, for the cell presented in Figure 3, the first figure received rank I, the second rank II, the fourth rank III and the third rank IV. When the figure appeared simultaneously with a background, the same ranking was also applied, so that a rank for a particular neuron always corresponded to the same figure whatever the background. Backgrounds were divided into two categories, according to the net response of the cell to the combination preferred figure–background. The background for which this combination gave the strongest response was assigned to the first category, i.e. background I, while the other background, eliciting the weaker response when combined with the preferred shape, is background II, a convention used throughout the remainder of this paper. For a given neuron, the same categorization of backgrounds was applied to the six other combinations of figures and backgrounds. Responses of each neuron were normalized with respect to the strongest response of the neuron amongst the 16 conditions and the normalized responses were averaged across cells.

In each of the four samples of neurons, the slope of the relationship between figure rank and average normalized response decreased sharply when a background was added (filled symbols). Adding the background resulted in a 2- to 5-fold reduction of the average response to the preferred figure. The effect of adding the background upon shape selectivity was tested on the net response by ANOVA and found to be significant in each of the four cases (interaction background and figure; $P < 0.001$ for each sample), indicating that at least the degree of shape selectivity was altered by adding the background shape.

Figure 5*A* plots the distribution of the response index for the shape-selective and for the shape- and color-selective cells. The distributions are clearly shifted towards positive responses indices (median 0.60), indicating that the response usually decreased when the background was added. A greater response for the figure–background combination than for the single figure (negative RI) was infrequently observed. The RI values for the two tasks were pooled since the degree of response reduction was similar in either behavioral paradigm: the median RI for the fixation task was 0.61 (n = 54) and 0.60 (n = 47) for the discrimination task [not significant (Mann–Whitney U-test; U = 1240; NS; n_1 = 54, n_2 = 47)), indicating that even when the subject is actively discriminating the figures and ignoring the background, these TE neurons are strongly affected by the presence of the background shape.

A distinction in the degree of reduction in response was observed between the color-selective units that were inhibited by color 2 and those that were not inhibited by the second color. Fifteen out of 46 cells (23%) were inhibited by the shapes in color 2. The median value of the response index was 1.13 for cells inhibited by color 2 and 0.59 for the other color-selective cells, a value not significantly different from the median RI of cells that were not color selective (0.50) (Mann–Whitney U-test; U = 727; NS; n_1 = 55, n_2 = 31) . The difference in RIs between cells inhibited by color 2 and those not inhibited by color 2 proved statistically significant (Mann–Whitney U-test; U = 136; $P < 0.05$). Thus it follows that, as long as the neuron is not inhibited by the background color, the degree of difference between the response to the color of figure and the color of the background does not affect the degree to which the response is modulated by the figure–background combination. The greater response attenuation with the background observed for cells inhibited by color 2 might be a direct result of the inhibitory effect of this color on the response of the neuron.

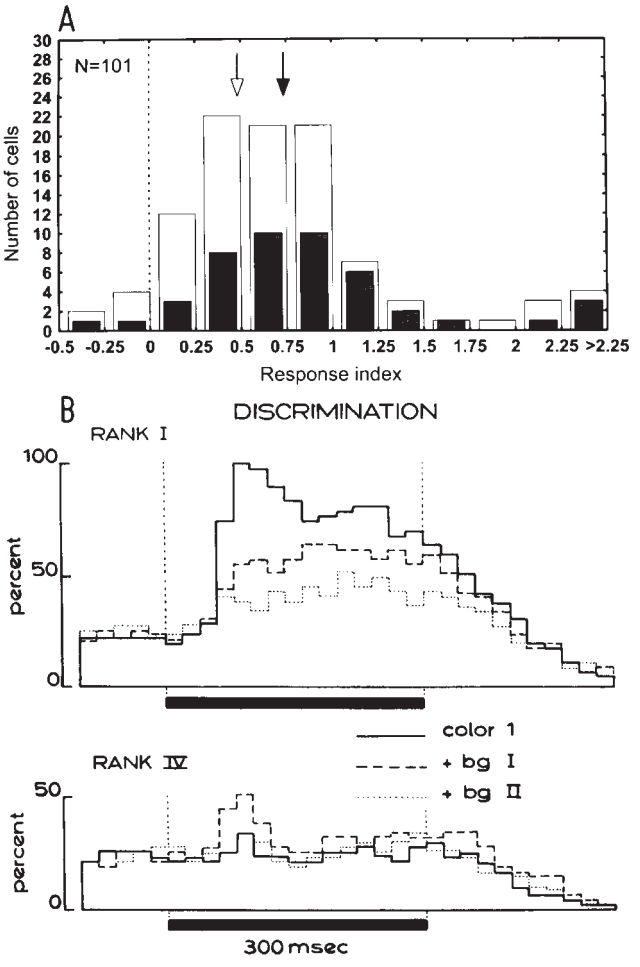

Figure 5. Distribution of the response index and population response. (*A*) Distribution of the response index for shape-selective cells (open bars) and shape- and color-selective cells (filled bars). Positive values represent units with suppression of the response in the conditions with background. Negative values represents units with an increase of the response in the same conditions. The arrows indicate the median of the distributions for shapes selective cells (open arrow) and shape and color selective cells (filled arrow). (*B*) Population peristimulus time histogram for the figure alone and figure + background conditions. Averaged normalized activity of all the cells recorded in the discrimination task (n = 47). For each unit the responses were normalized with respect to its maximum activity in a bin of 20 ms. Histograms of normalized activity were then averaged across neurons. The first row shows the response to the single preferred figure (rank I) in color 1 or to figure + background (+ bgI, + bgII). The second row shows the response in rank IV (least preferred figure) in the same conditions. The horizontal bar below each histogram indicates the time of stimulus presentation.

It is possible that a strong shape-selective response occurred even to the figure–background combination, but at a longer latency. To address this, we computed a population peristimulus time histogram by pooling the normalized activity of the neurons recorded during discrimination in bins of 20 ms. We pooled shape-selective and shape- and color-selective cells, since both types showed a reduction of the response for the figure–background combination. The population histograms for preferred figures (rank I) and figures of rank IV, when presented

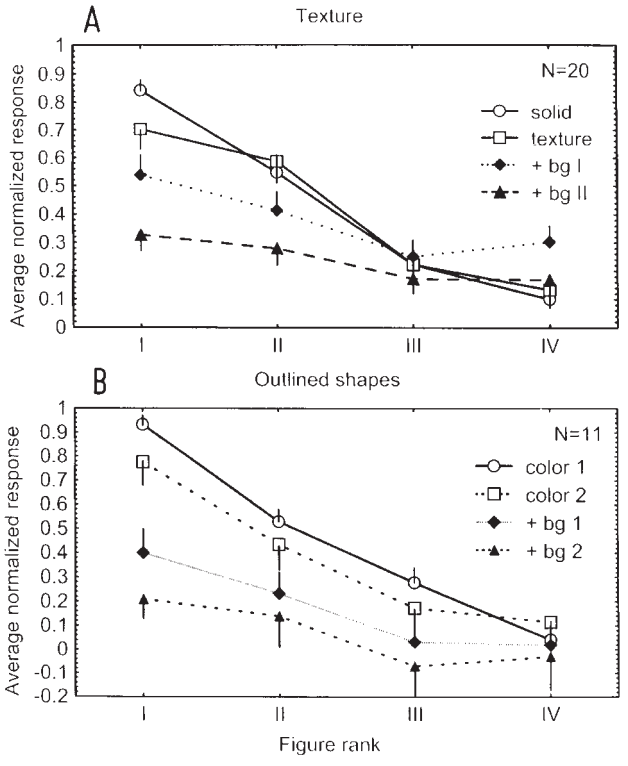

Figure 6. Effect of texture and outlines. (*A*) Texture as a segmenting cue. Average normalized response as a function of figure rank for single solid figures (solid line, open circles), for single textured figures (solid line, open squares) and solid figures + textured backgrounds (bgI, dotted line, filled diamonds; bgII, interrupted line, filled triangles). Twenty cells tested. The figure rank was determined using the response to single solid figures. (*B*) Effect of backgrounds for outlined shapes. Average normalized response as a function of figure rank for single outline figures in color 1 (solid line, open circles), in color 2 (stippled line, open squares) and outline figures + outline backgrounds (bg1, dotted line, filled diamonds; bg2, interrupted line, filled triangles). Eleven neurons tested. The figure rank was determined using the response to single outline figures in color 1. Normalization was done on each unit separately with respect to the strongest response of the cell.

alone (full line) or with either one of the two backgrounds (stippled), are shown in Figure 5*B*. The onset times of the responses to single figures and those of residual responses when a background was added were similar. It follows that the decrease in the average response observed when backgrounds were added cannot be attributed to a shift in the latency of the response under these conditions but represents a genuine response reduction.

Since the effect of a background was similar for cells recorded during the fixation task and the discrimination task, results were pooled in the study of the influence of the segmenting cues and of background shape properties.

Influence of Different Segmenting Cues

In order to test the influence of texture as segmenting cue, solid figures in a single color were presented alone or together with a textured background shape of the same color. The texture of the background shape was composed of a random dot pattern of 50% density (Fig. 1). Figure 6*A* shows the averaged normalized response as a function of figure rank for 20 cells tested with single figures alone (solid or textured) and solid figures combined with a textured background. Note that the response and shape selectivity are on average similar for solid or textured single figures (compare open circles to open squares). The presence of the textured background again reduced the average response level and altered the degree of shape selectivity [$F(6,114) = 12.59$; $P < 0.001$]. Also, the degree of response reduction was very similar to that obtained when figure and background differed in color: the median RI when texture was the segmenting cue was 0.56 (quartiles, 0.19–0.74; $n = 20$).

Outline figures (see Fig. 1) and outline background shapes were used to test whether the same effect could be generalized to unfilled shapes. Figure 6*B* shows the average normalized response as a function of figure rank for single outline figures and for the combination with outline backgrounds (11 cells tested, six of which were tested with two backgrounds; the responses to the figure–background combination of those cells tested with only one background are pooled with the responses to background 1 of the cells tested with two backgrounds and labelled background 1 in Fig. 6*B*). Nine of these cells were not color selective and showed an average RI of 0.59 (SE 0.14), while the response of the other two color selective units was also strongly reduced when the overlapping shape outlines were presented (mean RI: 0.68; SE 0.07). Note that in the case of outlines, the achromatic contrast of the figure with its immediate background is the same in the figure-alone and the overlapping-shapes conditions. The observed response reduction for the outlines strongly suggests that the response decrement in the case of the filled figure–background combinations is not merely due to a change in the local achromatic contrast of figure and background. This is supported by a within-cell comparison of the effect of the figure–background combination for filled and outlined shapes: the median RI was 0.48 (quartiles: 0.25–0.59; $n = 14$) for filled shapes and 0.54 (0.22–0.65; $n = 14$) for outlined shapes (Wilcoxon matched pairs test; $T = 43$; NS).

Influence of Background Stimulus Properties

Size of the Background Stimulus

Nine cells were tested with backgrounds of two different sizes (standard size and 20% size increase). Fifteen backgrounds were tested in total, three cells being tested with only one instead of two background shapes. None of the nine cells tested were color selective. The median response index for standard sized backgrounds was 0.23 (quartiles, –0.03–0.69; $n = 15$) and 0.59 (quartiles, 0.15–0.92; $n = 15$) for the larger ones, a difference which proved statistically significant (Wilcoxon matched pairs test; $T = 11$; $P < 0.005$; $n = 15$). Thus, the larger the background shape, the stronger the response reduction, at least for neurons that are not that strongly suppressed by the smaller background.

Shape of the Background Stimulus

We tested 14 cells with at least three, usually four different backgrounds. Five of these 14 cells were shape and color selective but only one cell was inhibited by color 2. Figure 7*A* shows the response of one neuron to four figure–background combinations ranked according to the net response of the cell to the single figures. The shape selective response was strongly and differentially affected by the different backgrounds used [interaction rank- background significant: ANOVA, $F(9,144) = 3.18$; $P < 0.001$]: background 1 had overall a weak effect whereas background 3 changed the selective response completely. Figure 7*B* shows a cell whose response changed weakly with background 4 but was very strongly affected by backgrounds 1–3. Again, the interaction between figure rank and background

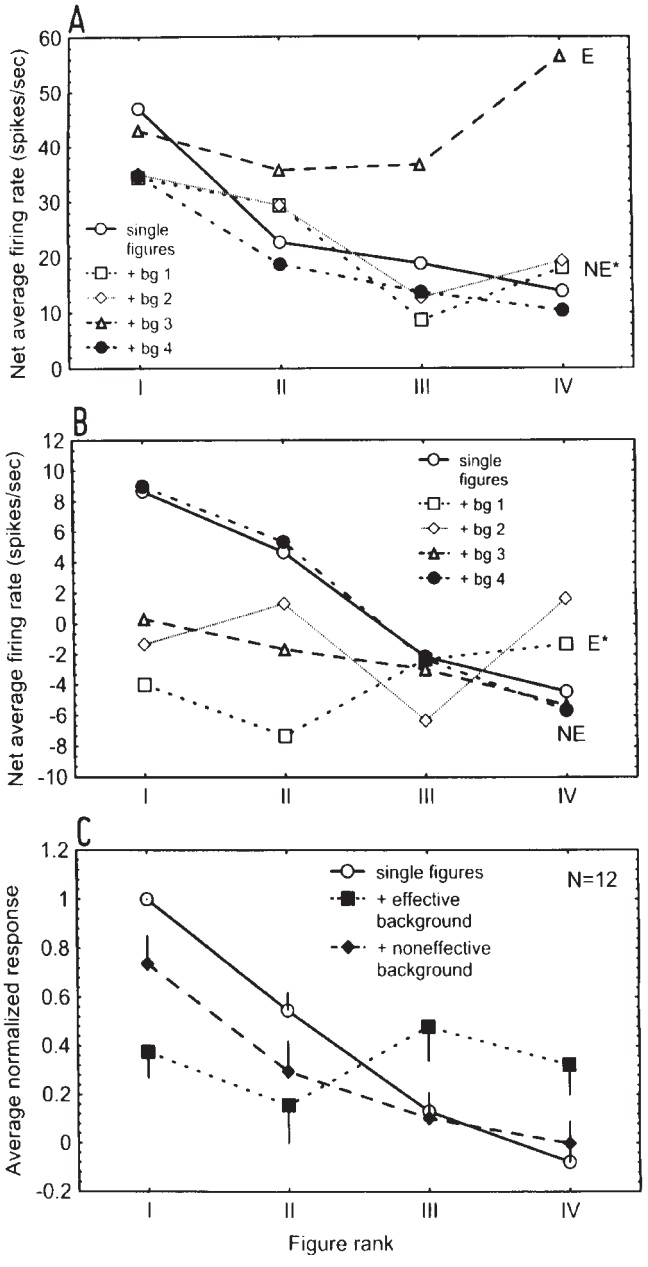

Figure 7. Influence of the shape of the background stimulus. Response strength as a function of figure rank for two single cells (*A*, *B*) tested with four different background shapes (bg1, bg2, bg3, bg4). The response to single figures (solid line, open circles) shows the selectivity of the cell and served to rank the figures and the different combinations figure + background. Effective backgrounds are indicated by E, non-effective ones by NE. A star (*) indicates which background, E or NE, had the largest size. (*C*) Average of 12 neurons. Average normalized response as a function of figure rank for single figures in color 1 (solid line, open circles), figures + effective backgrounds (dotted line, filled squares) and figures + uneffective backgrounds (interrupted line, filled diamonds). The figure rank was determined using the response to single outline figures in color 1. Normalization was done on each unit separately with respect to the response of the cell to the single figure in rank I.

shape was statistically significant [ANOVA, $F(9,144) = 2.27$; $P < 0.02$].

In order to show the differential effect of the background shape, two shapes were selected for each cell. 'Noneffective' backgrounds for all four figures, i.e. those that changed the selectivity of the cells the least, were put into one category (e.g. backgrounds labeled NE in Fig. 7*A*); the second category contained 'effective' backgrounds, those that affected shape selectivity the most (e.g. backgrounds labeled E in Fig. 7*A*). An average of the response with and without these two selected backgrounds is given in Figure 7*C* for the 14 neurons. The curve for the combination figures + noneffective backgrounds (filled diamonds) shows the same correlation between response and figure rank as the curve for single figures. Shape selectivity was unaltered for these backgrounds. The curve for figures + effective backgrounds shows no correlation between rank and response, indicating that shape selectivity is strongly modified. The interaction between figure rank and background type (effective vs noneffective) was statistically significant [ANOVA, $F(3,39) = 7.55$; $P < 0.001$; $n = 14$]. We compared the size of effective backgrounds to the size of ineffective ones in order to find out if a correlation exists between effectiveness and size. For the 14 cells tested, only six effective backgrounds were larger than the ineffective ones, and the median size of effective backgrounds (4.8; 3.2–5.4; $n = 14$) and the median size of noneffective ones (5.0;4.0–5.4; $n = 14$) were very similar. Thus, it can be concluded that there is a genuine influence of the shape of the background that is independent of its size.

Do TE Neurons Segment Figure from Background Shape?

One possible explanation of the effect of the background is that these TE neurons are unable to segment the figure from the differently colored or textured background shape. Thus, the figure–background combination may be treated by the neurons as a completely different stimulus and the altered response may merely be a result of the neuron responding differently to this altered stimulus. Adding differently shaped backgrounds to differently shaped figures results in a diversity of combined shapes, and thus a neuron may respond to some of these figure–background combinations but not to others, explaining the above-described influence of background shape on the apparent suppression of the response.

However, for some neurons and background shapes there was still a response to the figure–background combination, and on average the shape preference of this residual response was similar to that obtained with the figures alone (see Figs 4 and 7), apparently suggesting some segmentation of figure and background. Still, residual responses to the figure–background combination may not be due to segmentation but merely reflect a response to a feature common to both the figure and the figure–background combination. This was tested by changing the color inside the figure–background stimulus: neurons were tested with identical figure–background combinations for which (i) the color of figure and background differed (segmentation condition); (ii) the color was the same for figure and background (no-segmentation condition; see Fig. 1); or (iii) upper and lower part of the combined stimulus differed in color [altered segmentation condition; see Fig. 1(cut)]. If the residual response of the cell is due to segmentation of figure and background shape, the response should be different for the segmentation and the two other conditions, since the figure is no longer (perceptually) present in the latter case. However, if the residual response is due to some feature common to the figure and the figure–background combination (e.g. common in the contour of both types of stimuli) but unrelated to the segmentation and coloration of the figure and background, than the responses should be similar in the three conditions. For each neuron we selected background 1, which was combined with the same figures as those used in the standard test. Figure 8 shows the

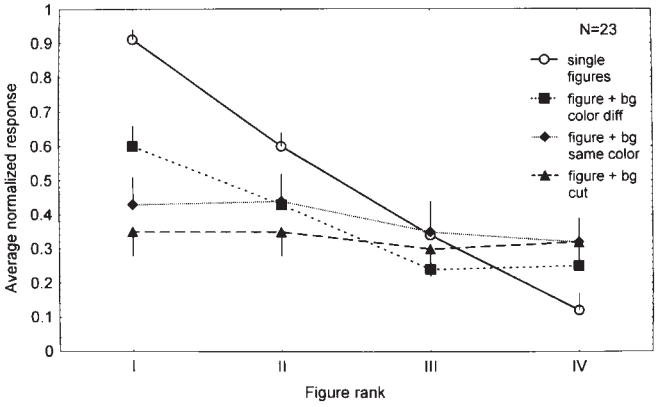

Figure 8. Response to figure–background stimuli with different kinds of segmentation. Average normalized response as a function of figure rank for singles figures (solid line, open circles), figures + background in different colors (segmentation of figure and background; dashed line, filled squares), figures + background in the same color (no segmentation; dotted line, filled diamonds) and figures + background 'cut' (altered segmentation; interrupted line, filled triangles). The figure rank was determined using the response to single figures in color 1. Normalization was done on each unit separately ($n = 23$) with respect to the strongest response of the cell.

average response of 23 neurons tested in the fixation paradigm. Figures were ranked according to the net response to the best figure presented in isolation in the best color and the same rank was applied for all other conditions. It can be seen that the averaged normalized response to the standard figure–background combination is reduced but still decreases with increasing figure rank, as it does for the single figures. When figure and background had the same color (no segmentation) and when 'cut stimuli' (altered segmentation) were used, a more dramatic effect was observed: the average response was independent of figure rank. The interaction between figure rank and background condition (segmentation, no segmentation and altered segmentation) was statistically significant [ANOVA; $F(6,132) = 3.80$; $P < 0.001$; $n = 23$]. Thus, these results suggest that the residual response of the cells was due to the figure that was segmented from the background by virtue of color differences between figure and background.

Discussion

We found that the response and selectivity of IT neurons to a given filled or outline shape (figure) can be dramatically altered when this shape overlaps another shape (background). This response alteration is present when both shapes are perceptually segmentable and even when the subject actively discriminates the figures against varying backgrounds. The interaction between the two shapes is dependent on figure and background shape. The most commonly found response alteration was a strong reduction of the response to the figure in the presence of the background shape, although more complex interactions of figure and background shape were also observed.

In the discrimination paradigm, in which the monkey segments and presumably attends the figure, strong shape interactions were observed. This unexpected result does not contradict earlier studies on attention effects in IT. In these studies, attention was directed either to a particular, cued location in space (Moran and Desimone, 1985) or to a particular, cued target (Chelazzi *et al.*, 1993). Thus, a location or target was actively held in working memory and could bias the sensitivity of the neurons in a location- or target-specific way [e.g. suppression of the response for cells at unattended locations or to unattended targets (see Desimone and Duncan, 1995)]. In our case such a top-down influence is less likely to have occurred since the subject was not cued for a particular target: he was required to discriminate 10 shapes which were presented in an unpredictable, random order and which could overlap another, irrelevant shape. Thus, the behavior in our task was more bottom-up driven, and our results show that under these conditions strong figure–background interactions are present in IT cortex.

The surprisingly strong effect of the background on the responses to the figure can be interpreted in different ways. Firstly, it is possible that the neuron does not segregate figure from background shapes and responds to the figure–background combination as just another shape, unrelated to the figure. In this case, the reduction of the response for the figure–background combination is simply due to the neuron not being responsive to that stimulus. An alternative interpretation is that the response alterations result from interactions between figure and background stimulus, reminiscent of the interactions between simultaneously presented patterns of different orientation or direction of motion observed in V1 (Knierim and Van Essen, 1992; Sillito *et al.*, 1995) and in areas MT (Allman *et al.*, 1985; Tanaka *et al.*, 1986; Raiguel *et al.*, 1995) and MST (Tanaka *et al.*, 1986; Lagae *et al.*, 1994). Our data favor the latter interpretation since for some background shapes the average response to the combination of the background and preferred figure was larger than the response when the unpreferred figures were combined with the background. The results of our control test, in which the segmentation cue was absent or altered, showed that this remaining selectivity was not due to the shape of the composite itself (see Fig. 8). Thus, the response of the neuron is partially determined by the figure, despite being embedded in the background shape.

It should be noted that color is a powerful segmentation cue [for evidence of preattentive, parallel color segmentation see Nothdurft (1993)], and that figure and background also differed in luminance in addition to a color or texture difference. Thus, the suppression is unlikely to be due to the use of these particular segmentation cues. However, additional cues such as motion contrast (Sary *et al.*, 1993) or binocular disparity may modulate the suppression, as has been observed for motion transparancy displays in area MT (Bradley *et al.*, 1995).

Since the single filled figures were presented on an uniform black background and the filled figures overlapped the background shape, the contrast of the filled figure with its immediate background was reduced – which might have reduced the magnitude of the response to the overlapping shapes. However, the presence of the response reduction for filled and outlined shapes in color selective units, which could differentiate figure and background by virtue of their color contrast, as well as the effect of different background shapes having the same figure–background contrast, suggest that the response reduction for the filled figures does not merely result from this achromatic contrast reduction. Also, the presence of a strong suppression for overlapping outline shapes, which are not affected by this contrast problem and can be segmented even when they do not differ in color by virtue of local form cues (e.g. X-junctions and Gestalt law of continuation), shows that the suppressive effect when adding two shapes is not a simple luminance-interaction effect.

That other shapes simultaneously present in the immediate neighborhood of the preferred shape strongly affects the

response of a TE neuron does not disagree with results obtained with the reduction method used by Tanaka and co-workers (Tanaka *et al.*, 1991; for review see Tanaka, 1996). In the reduction method one starts with a complex stimulus to which the cell responds and systematically strips out components of the stimulus as long as the response remains unaffected. Given our findings, using the reverse approach of making an effective stimulus more complex, this would mean that the other features present in the complex stimulus were not effective in altering the response to the critical feature, as was the case with some of the backgrounds in our study. Our results show that the response of TE neurons to a complex stimulus cannot always be predicted from the responses to components of the stimulus, as has been noted by Albright and Gross (1990).

It has been reported previously (Sato, 1988, 1989; Miller *et al.*, 1993; Rolls and Tovee, 1995) that the simultaneous presence of two non-overlapping stimuli elicit less response from IT neurons on average than the presentation of either shape alone. In those studies, the response to two *non*-overlapping stimuli was on average ~75% of the response to a single stimulus, while the response to the overlapping shapes in our study was usually only 30% of the response to the isolated figure. A comparison of these studies suggests that overlapping shapes produce stronger reduction of the response than non-overlapping stimuli. Overlapping shape parts can be interpreted as belonging to the same object, especially when they are presented at the same depth, whereas two disjoined shapes can belong to different objects. The response suppression we observed suggests that overlapping shapes are treated as *single* objects by TE units and that their response not only depends on the presence of their preferred shape but also on the configuration of which this shape is a part. Indeed, many objects consist of particular combinations of particular shape parts. Thus, suppressed responses to combinations of particular shape parts, while allowing a response to combinations of other parts corresponding to those in a real object, will increase the selectivity of a system coding for these objects.

In general, the present results show that it is difficult to predict the response of a neuron to a compound stimulus obtained by adding a second stimulus to an existing stimulus, if only the response of that neuron to the latter stimulus alone is known. The significance of this conclusion can be appreciated in light of the fact that our visual world rarely consists of an isolated, simple stimulus.

Notes

This work was supported by FWO-Vlaanderen grant G. 0172.96, IUAP 22, G.O.A. 95-6, GSKE. M.M. was on leave from and partially supported by the Laboratory of Neurophysiology, Université Catholique de Louvain. R.V. is a research associate of FWO-Vlaanderen. We thank M. DePaep, G. Vanparijs, P. Kayenberg and G. Meulemans for technical assistance, W. Spileers for surgical assistance, L. Arckens for histological assistance and S. Raiguel for critical reading of the manuscript. Chao-Yi Li participated in some of the experiments.

Address correspondence to Rufin Vogels, Laboratorium voor Neuro-en Psychofysiologie, Faculteit Geneeskunde, KULeuven, Campus Gasthuisberg, Herestraat, B-3000 Leuven, Belgium.

References

Albright TD, Gross CG (1990) Do inferior temporal cortex neurons encode shape by acting as Fourier Descriptor filters? In: Proceedings of the International Conference on Fuzzy Logic and Neural Networks, Izuka, Japan, pp 375–378.

Allman J, Miezin F, McGuinness E (1985) Stimulus specific responses from beyond the classical receptive field. Annu Rev Neurosci 8:407–430.

Bradley, DC, Qian, N, Andersen, RA (1995) Integration of motion and stereopsis in middle temporal cortical area of macaques. Nature 373:609–611.

Chelazzi L, Miller EK, Duncan J, Desimone R (1993) A neural basis of visual search in inferior temporal cortex. Nature 363:345–347.

Dean P (1976) Effects of inferotemporal lesions on the behavior of monkeys. Psychol Bull 83:41–71.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

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

Honig MG and Hume RI (1989) Dil and DiO: versatile fluorescent dyes for neuronal labelling and pathway tracing. Trends Neurosci 12:333–341.

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

Kirk RE (1968) Experimental design: procedure for the behavioral sciences. Belmont, CA: Brooks-Cole.

Knierim JJ, Van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 67:961–980.

Komatsu H, Ideura Y, Kaji S, Yamane S (1992) Color selectivity in the inferior temporal cortex of the awake macaque monkey. J Neurosci 12:408–424.

Kovacs G, Vogels R, Orban GA (1995) Selectivity of macaque inferior temporal neurons for partially occluded shapes. J Neurosci 15:1984–1997.

Lagae L, Maes H, Raiguel S, Xiao DK, Orban GA (1994) Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. J Neurophysiol 71:1597–1626.

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

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

Missal M, Vogels R, Orban GA (1995) Neuronal responses in inferior temporal cortex and figure-ground segmentation. Soc Neurosci Abstr 21:19.

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

Nothdurft, HC (1993) The role of features in preattentive vision: comparison of orientation, motion and color cues. Vision Res 14:1937–1958.

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) The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. Exp Brain Res 103:409–420.

Raiguel S, VanHulle MM, Xiao DK, Marcar VL, Orban GA (1995) Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. Eur J Neurosci 7:2064–2082.

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

Sary Gy, Vogels R, Kovacs Gy, Orban GA (1995) Responses of inferior temporal neurons to luminance-, motion, and texture-defined gratings. J Neurophysiol 73: 1341–1354.

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

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

Schwartz EL, Desimone R, Albright TD, Gross CG (1983) Shape recognition and inferior temporal neurons. Proc Natl Acad Sci USA 80:5776–5778.

Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J (1995) Visual cortical mechanisms detecting focal orientation discontinuities. Nature 378:492–496.

Snodderly DM and Gur M (1995) Organisation of striate cortex of alert, trained monkeys (*Macaca fascicularis*): ongoing activity, stimulus

selectivity, and widths of receptive field activating regions. J Neurophysiol 74:2100–2125.
Tanaka K (1996) Inferotemporal cortex and object vision. Annu Rev Neurosci 19:109–139.
Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of local and wide field movements in the superior temporal visual areas of the macaque monkey. J Neurosci 6:134–144.
Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol 66:170–189.
Vogels R, Orban GA (1996) Coding of stimulus invariances by inferior temporal neurons. Progress Brain Res 112:195–211.