

Shape-coding in IT cells generalizes over contrast and mirror reversal, but not figure-ground reversal

Gordon C. Baylis^{1,2} and Jon Driver³

¹ University of Plymouth, Plymouth Institute of Neuroscience, 12 Kirkby Place, Plymouth, PL4 8AA, UK

² Department of Psychology, University of South Carolina, Columbia, South Carolina 29208, USA

³ Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

Correspondence should be addressed to G.C.B. (gordon@pion.ac.uk) or J.D. (j.driver@ucl.ac.uk)

We assessed how the visual shape preferences of neurons in the inferior temporal cortex of awake, behaving monkeys generalized across three different stimulus transformations. Stimulus-preferences of particular cells among different polygon displays were correlated across reversed contrast polarity or mirror reversal, but not across figure-ground reversal. This corresponds with psychological findings on human shape judgments. Our results imply that neurons in inferior temporal cortex respond to components of visual shape derived only after figure-ground assignment of contours, not to the contours themselves.

Inferotemporal (IT) cortex is involved in visual shape representation and visual object recognition, based on evidence from single-cell recording^{1–9}, functional imaging^{10,11} and lesion studies^{9,12}. In comparison with earlier visual areas, cells in IT have larger receptive fields and show more abstract preferences for complex shape properties^{2,4,6,9,13}, but exactly how this region represents shape remains controversial^{3–5,13}. Here we examined shape representation within IT in relation to figure-ground reversal, as well as other stimulus manipulations that served as control comparisons.

The figure-ground assignment of a given visual display can dramatically alter the shape that human observers perceive (examples, top of Fig. 1). Adjacent figure and ground regions defined by a common contour are perceived as very different. Human observers typically recognize the figure later (for example, the face in the top row of Fig. 1), but not the ground (white shape in that row), even for judgments based on exactly the same shared contour^{14–18}. Moreover, they rate a mirror image of the figure as more similar¹⁹ to the original figure-ground display than an image of the ground in isolation. This arises even though the ground probe shares exactly the same curved contour as in the originally exposed display, whereas the mirror image of the figure has a mirror-reversed contour. These phenomena also arise for shapes made by unfamiliar contours^{15–20} (see below), not just for profiles of meaningful shapes. Such effects reveal the influence of one-sided edge assignment on visual shape perception in humans^{15–20}.

Here we examined how the shape preferences of IT cells in the primate brain may relate to these psychological phenomena. Specifically, we tested how the preferences of individual IT cells for stimuli drawn from a population of pseudorandom two-dimensional shapes would generalize across three different transformations: figure-ground reversal, reversal of contrast-polarity and mirror-image reflection about the vertical (Fig. 1a–h). All shapes were polygons with straight edges at the top, bottom and along one side, and with a pseudo-randomly curved contour on

the other side^{15–20}. The curved contours of possible polygons differed in their identity and location (left or right side of polygon).

It is possible that any selectivity in the responses of IT cells to these stimuli is determined just by these physical differences among the displays. Alternatively, IT responses might show patterns that are more like shape judgments in human observers, where figure and ground regions are perceived to have very different shapes despite their common defining contour, with the mirror image of any figure being perceived as more similar to that figure than its ground (as confirmed for the present stimuli also; see below). For the displays used here, exactly the same curved contour was present across a reversal of figure-ground assignment (Fig. 1, compare a to e, c to g, b to f, and d to h), yet this contour produces shapes that look very different to human observers when figure and ground are reversed^{14–20}.

The curved contour was necessarily on opposite sides of the figure region versus the adjacent ground region within any display (stimuli a–h, Fig. 1). Our further manipulation of mirror-imaging (see also ref. 6) controlled for this, as the curved contour of any mirror image of an original figure is on the same side as the curved contour of the original ground (stimuli a–h, Fig. 1). The figure and ground region of each individual display also differed in contrast polarity (one white, the other black). Our orthogonal manipulation of reversing contrast-polarity (see also ref. 8) controlled for this, as a contrast-reversal of an original figure has the same polarity as the original ground (stimuli a–h, Fig. 1).

We recorded activity from IT cells in monkeys to determine their firing rates for the different stimuli, and to determine how these rates correlated across the three different stimulus transforms. We also required human observers to rate the similarity of the displays across the same transforms.

RESULTS

We recorded from 88 cells in areas TEa, TEm and TE3 (ref. 20) of 2 awake monkeys while they viewed displays drawn from 32

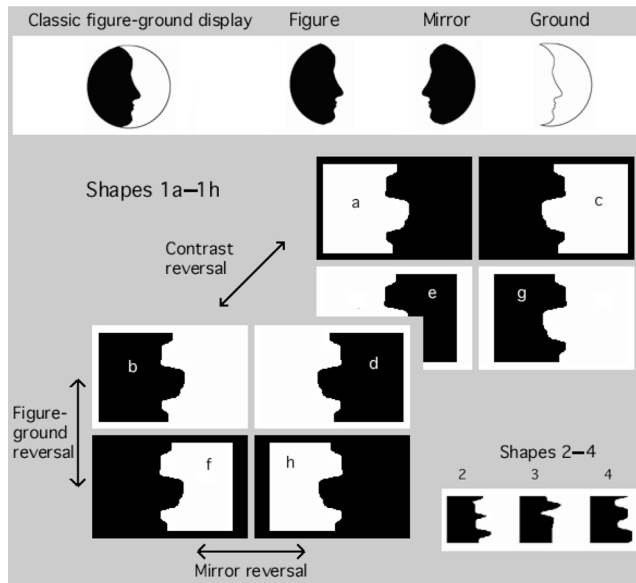


Fig. 1. Example stimuli. Top, classic figure-ground display, together with its components. Humans rate a mirror image of the figure as more similar to the original figure-ground display than the original ground in isolation. Stimuli a-f, Visual displays for the single-cell recording experiment, showing how 8 different displays were generated from one particular curved contour (type 1 is shown). Bottom right, 3 additional types of curved contours (2-4); each of these analogously generated a 8 different displays (2a-h, 3a-h, 4a-h). Three aspects of the displays were manipulated orthogonally, in a $2 \times 2 \times 2$ fashion illustrated by the layout of shapes 1a-h, which shows one 2×2 table of possible displays in the 'front' plane (b, d, f, h), with another 2×2 table of possible displays in the 'back' plane (a, c, e, g). All displays comprised either a white filled polygon on a black background, or vice versa. The difference between examples in the front plane and the back plane in the illustration depicts this contrast-polarity transform for otherwise equivalent displays. Each display also appeared in mirror image form. The difference between examples in adjacent columns for each of the 2×2 tables in the schematic illustrates this mirror-reversal transform. Finally, a given curved contour could have the figural region (as defined by surroundedness^{14,17}) on its left or right, leading to the figure-ground transform between examples in adjacent rows for each 2×2 table in the schematic. Only the figure-ground transform leaves the curved contour entirely unchanged.

possibilities (stimuli a-h, Fig. 1, equivalent transforms were implemented for types 2, 3 and 4, thus yielding $8 \times 4 = 32$ stimuli in total). Eighty-nine percent of cells (78/88) showed significant differences in mean evoked firing rate in the interval 100 to 600 ms after stimulus onset, as a function of which of the 32 possible stimuli were shown (at $p < 0.01$ or better). We then examined how the shape preferences revealed by these differential evoked firing rates correlated across the three transformations we had applied to the stimuli. Most cells showed significant and substantial correlations in stimulus preferences across reversals of contrast polarity and across mirror imaging, but not across figure-ground reversal.

We first show the correlations across these three transforms for one illustrative neuron (Fig. 2). All 32 stimuli contributed to each of these correlations, but stimuli were paired differently for each correlation (see Methods). We plot the mean firing rates of this cell (in the 100-600 ms interval after stimulus onset) for all

32 stimuli (Fig. 3; the same nomenclature is used for these stimuli as in Fig. 1). The pattern of firing rates was similar across the contrast-reversal and mirror-image transforms, but not across the figure-ground reversal. (Compare the four appropriate pairings of graphs, each with four bars, across each of these transforms, Fig. 3.) Peri-stimulus time histograms for this cell in response to the 16 stimuli generated from types 1 and 2 show that responses were similar across mirror-image and especially contrast-reversal transforms, but that they differed markedly across the figure-ground transform (Fig. 4). For instance, type 2 received a much more vigorous response than type 1 in version b, but the opposite ordering applied for version f (the figure-ground transform of version b).

In histograms of the distributions of correlations for all cells in the population across the three transformations, most cells showed significant positive correlations in shape preference across reversed contrast polarity in the display (mean correlation coef-

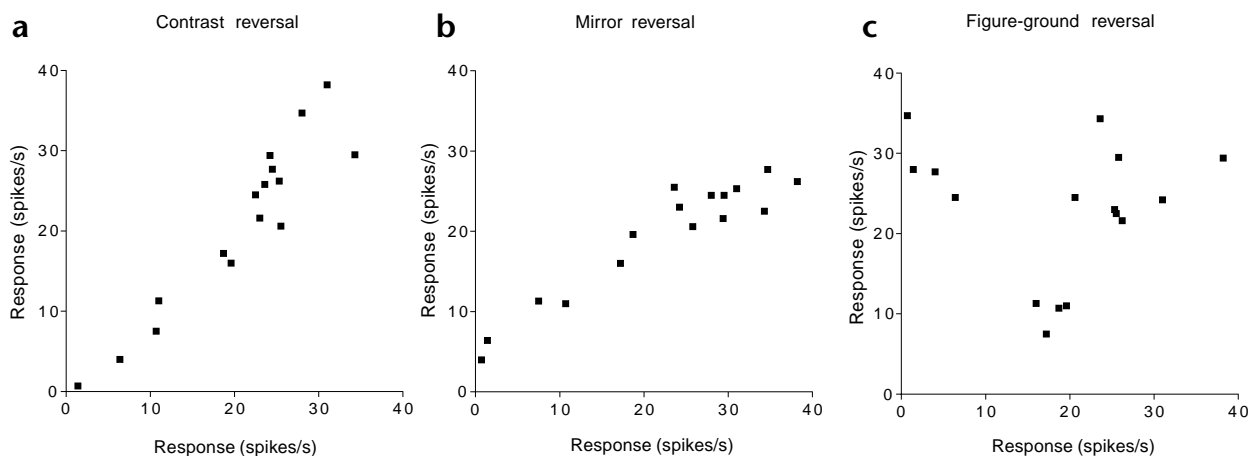


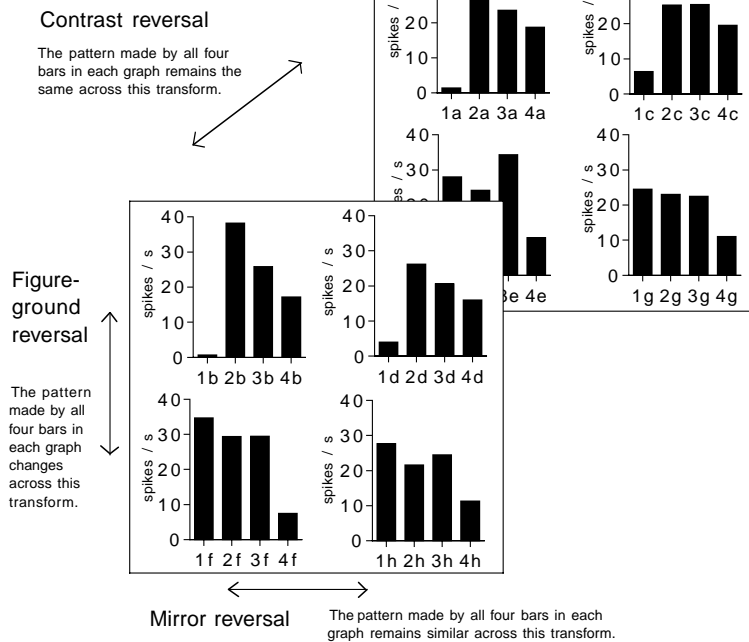
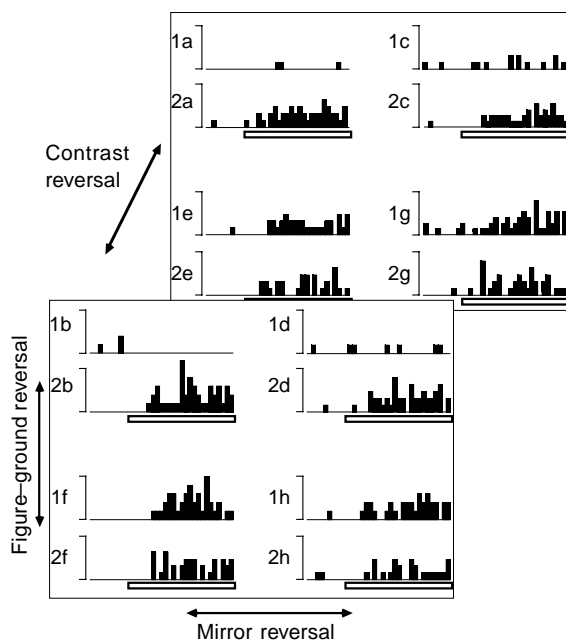
Fig. 2. Correlation plots for a single cell. Plots show mean firing rates (in the period 100-600 ms after the stimulus) for an illustrative neuron, for particular stimuli along the x-axis, and for transformed versions of the same stimuli along the y-axis. (a) Contrast reversal transform. (b) Mirror reversal. (c) Figure-ground reversal. The total set of 32 stimuli all contribute to each plot. For each plot, this set was divided into 2 subsets of 16, with each member of one subset providing a transformed version for one member of the other subset. (a, b) Stimuli that induced a particular firing rate led to a similar rate when transformed (correlations of 0.92 and 0.68 respectively, for this particular cell). No such relationship is apparent in (c) across the figure-ground transform ($R = 0.0$).



Fig. 3. Firing rates in response to the 32 stimuli for a single cell. Histograms show mean firing rate in the 100–600 ms period after stimulus onset for the illustrative cell from Fig. 2, now shown for each individual stimulus. The layout in the illustration has the same $2 \times 2 \times 2$ arrangement as in the schematic in the center of Fig. 1, and uses the same nomenclature for the 32 different stimuli (variants a–h on contour types 1–4). Hence, comparing laterally adjacent pairs of histograms addresses the mirror image transform of the stimuli; comparing histograms between the apparent ‘front’ and ‘back’ plane addresses contrast reversal; vertically adjacent histograms represent a figure–ground reversal. The pattern seen within each of the paired histograms stays similar across both the contrast and mirror-image transforms, but not across the figure–ground reversals, hence the correlations in Fig. 2 for the same cell.

ficient, $R = 0.59$), and likewise across mirror-image reflection of the presented shape (mean $R = 0.46$; Fig. 5). In contrast, correlation coefficients for figure–ground reversal were typically low, and centered around zero (mean $R = 0.04$). Chi-square tests, comparing the number of cells showing significant correlations across the different transforms, found many more such correlations for contrast versus figure–ground reversal ($\chi^2_1 = 80.4$, $p < 0.0001$), and for mirror-image versus figure–ground reversal ($\chi^2_1 = 44.96$, $p < 0.0001$). In addition, correlations were somewhat more pronounced for contrast than mirror image reversal ($\chi^2_1 = 9.44$, $p < 0.05$), in accord with the human similarity ratings reported below.

The poor generalization across figure–ground reversal was found equivalently for cells that showed a significant correlation across both contrast and mirror-image reversal (black, bottom histogram of Fig. 5) and those that did not (white, Fig. 5); these distributions for figure–ground reversal did not differ. We also assessed how the correlation coefficients developed as a function of time after stimulus onset. For contrast-polarity and mirror-image transformations, the average coefficients climbed rapidly, reaching asymptote at around 200–300 ms post stimulus onset.



For figure–ground reversals, the correlation coefficient averaged very close to zero throughout the trial.

Finally, we examined how the average firing rates changed during the trial for preferred versus non-preferred stimuli, and how this generalized across the three transformations. To do this, we first determined for each cell which of the 32 stimuli produced the maximal mean firing rate in a 100–600 ms time bin after stimulus onset; this defined the preferred stimulus for that cell. We also identified the non-preferred stimulus for each cell, producing the lowest mean firing rate in the same time window. We show the mean firing rates across all cells for their preferred versus non-preferred stimuli, at different times after stimulus onset (Fig. 6a). Firing rates for contrast-polarity and mirror-image reversals of these stimuli show how the preference was largely maintained across both these transformations (Fig. 6b and c). In contrast, the preference disappeared across the figure–ground transformation, consistent with our other findings (Fig. 6d). We confirmed the site of cellular recordings by histology (Fig. 7).

In a matching task (see Methods) on the same shapes as used in our physiological work, 12 human observers selected the untransformed figure on 87.5% of trials, the contrast-reversed version on 68.3% trials and the mirror-reversed version on 54.2% of trials. The latter two transforms were each selected significantly more often ($p < 0.01$) than the figure–ground reversal (only 19.7%). Moreover, the figure–ground reversal was not selected any more often than a shape with an entirely different contour (20.3%), and most selections for either of these two types arose when they were the only two alternatives presented (see Methods). These data con-

Fig. 4. Peri-stimulus time histograms of firing, using 20 ms bins, for the illustrative cell. Response to the 8 variants of type 1 and type 2 shapes. (Detailed responses to 16 stimuli are shown here, rather than all 32; type 3 and type 4 data (Fig. 3) are omitted for brevity.) $2 \times 2 \times 2$ layout and nomenclature for the stimuli are as in Figs. 2 and 3. The y-axis brace represents a firing rate of 100 spikes/s; the bar on the x-axis represents the first 500 ms of stimulus presentation time.

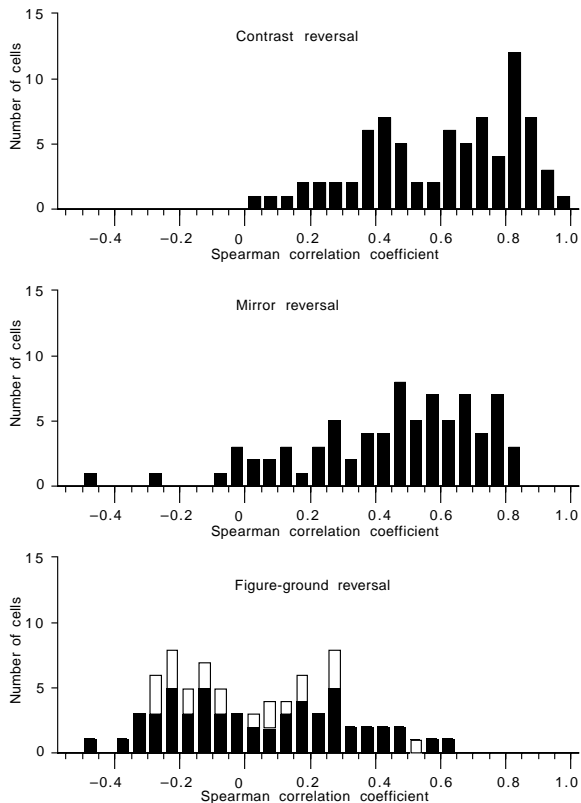


Fig. 5. Correlations of ranked stimulus preferences for each of the three transforms in the cell population. Histograms show the population distributions of Spearman rank-order correlations in firing rate (for 100–600 ms following stimulus onset) between transformed versions of the stimuli. Each bar indicates the number of cells from the population showing a particular size of correlation. Most cells show reliable positive correlations (with 15 degrees of freedom) across the contrast-reversal transform and mirror-reversal transform. Correlations for the figure-ground transform are much lower overall, averaging near zero. For figure-ground reversal plot, cells that showed significant correlations for both contrast and mirror-image reversal are represented in black; those that did not, in white.

firm that human perception of shape generalizes well across contrast reversal, and fairly well across mirror reversal. In contrast, figure-ground reversal alters shape perception as much as the generation of a new shape from an entirely different contour. Our findings for the shape preferences of IT cells in the primate brain closely parallel these aspects of human perception.

DISCUSSION

The shape preferences of IT cells generalized well across contrast reversals of the stimuli and across mirror imaging of the stimuli, but not across figure-ground reversals. The two-dimensional polygons we used varied in their particular curved contour. The other lines (three straight edges) were held constant in the stimulus sets assessed for each transformation, as we tested for generalization across contrast reversal, mirror imaging or figure-ground reversal. These three transformations have very dif-

ferent influences on the critical curved contour. Contrast reversal changes the polarity of this critical contour. Mirror reversal reflects this contour about the vertical. Only figure-ground reversal leaves the critical curved contour itself unchanged. (Its relative position with respect to the body of the shape changes, but this is applied equally to the mirror-reversal transform.) Thus, if the selective responses of IT cells had been caused primarily by just the curved contour that distinguished the various displays physically, then we should have found maximum generalization across figure-ground reversal, as only this keeps the curved contour constant. However, the opposite result was found, with generalization absent only for the figure-ground transform.

This demonstrates that the selectivity of IT responses is not determined simply by the distinctive contours in a display, contrary to simple edge-based models of shape recognition discussed elsewhere^{5,22}. Instead, coding in IT follows similar principles to that observed for human shape judgments. Human observers rate the mirror image of an original figure as more similar to the original display than the original ground^{18,19}, as confirmed here for the displays used in our physiological work. This arises even though the ground shares the same informative contour as the original figure, and hence has the ‘profile’ of the original figure embedded in it as background. We found here that IT cells likewise generalized more strongly across mirror imaging than across figure-ground reversal. Our findings for mirror imaging are con-

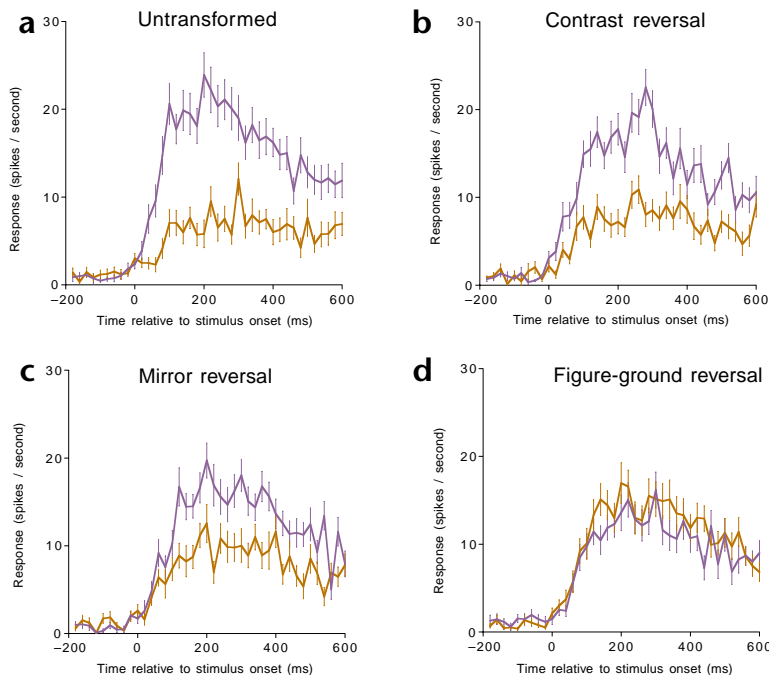


Fig. 6. Population response to preferred versus non-preferred stimuli across the three stimulus transforms. Mean firing rate across the population of neurons, for successive 20-ms time bins, with standard errors. (a) Responses to the preferred (blue) and non-preferred (red) stimulus selected for each neuron. (b) Responses to the contrast-reversed versions of each neuron’s preferred versus non-preferred stimuli, showing that the preference is maintained. (c) Responses to the mirror imaged versions of each neuron’s preferred versus non-preferred stimuli, again showing that the preference is still maintained. (d) Responses to the figure-ground reversed versions of each neuron’s preferred versus non-preferred stimuli, showing that the preference is now abolished.

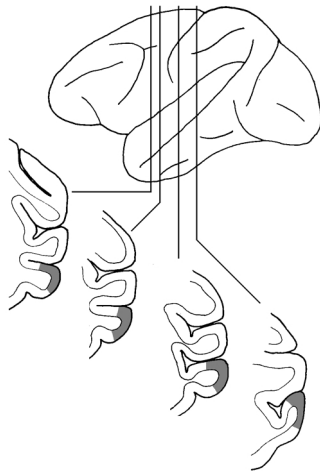


Fig. 7. Regions of IT cortex in which the cells were recorded, drawn on sections from the brain of monkey A. Top, locations of these coronal sections are shown on a schematic monkey brain.

sistent with other single-cell evidence⁶; the contrast-reversal findings also agree with previous studies^{8,23}. The additional comparison with figure-ground reversal here reveals that the selective responses of IT neurons correspond with psychological observations^{15–20} on how one-sided assignment of edges to figures constrains human perception of contoured shapes, and accord with human similarity ratings.

A longstanding question in vision research^{14,15,18–20} is why figures and their abutting grounds are perceived as so different in shape, despite the common contour. One computational proposal is that the visual system may decompose shapes into convex parts^{16,19,24}. A convexity in the outline of a figure (for example, the ‘nose’ in the face profile at the top left of Fig. 1) will produce a corresponding concavity in the abutting ground region of the image, and vice-versa¹⁹. This will lead to different convex parts on either side of a given contour. Our finding that IT neurons were driven by the figural shapes resulting from one-sided edge assignment, not by contours *per se*, seems consistent with shape representation within IT in terms of the layout of such component parts^{5,19,24}.

We re-analyzed the data in terms of another transform among the stimuli, to assess this hypothesis further. We correlated stimulus preferences across a transform that can be illustrated with reference to Fig. 1, comparing the lower left stimulus in the front panel (stimulus f, Fig. 1) to the top right stimulus in the back panel (stimulus c, Fig. 1), and so on. We thus compared pairs of stimuli that had the same contrast polarity and faced in the same direction, but had just the curved edge itself (not the shape as a whole) reflected. Thus, after figural assignment, the two members of the pair should have different convex parts. (Indeed, the change in convex parts is the same as the change for a figure-ground reversal, except that the parts now ‘point’ in the same direction.) We found that the (null) correlations across this transform were equivalent to those for our standard figure-ground reversal transform, averaging 0.09 versus 0.04, respectively, with no difference between the population distributions of these correlations.

Taken together, our results accord with theories of object recognition that propose^{15–20,25,26} that one-sided edge assignment precedes shape description in the visual system, with decomposition into component parts proceeding only for the figural side of any contour. A rival account²⁷ proposes instead that part decomposition initially arises for both sides of every

contour. The present study finds no support for the latter view at the level of IT responses, as the cells did not respond differentially to the presence of their preferred ‘profile’ in the background to the current stimulus (for example, Fig. 6d). Instead, our results show that shape description in IT cortex is entirely constrained by one-sided assignment of contours to figural objects.

METHODS

Animals and surgery. The experiment was conducted with two male macaque monkeys (*Macaca fascicularis*, 4.8 and 5.8 kg). With aseptic surgery, we placed a recording chamber and inserted a scleral coil in the left eye. All procedures were approved by the Institutional Animal Care and Use Committee.

Recording techniques. The activity of single neurons was recorded with epoxy-insulated tungsten microelectrodes (FHC, Brunswick, Maine) as the monkey sat in a primate chair, using standard techniques for single-cell recording⁷. Action potentials of single cells were amplified using BAK neurophysiological hardware, passed through a dual-window discriminator, with output TTL pulses timed to a resolution of 0.1 ms by the computer controlling the experiment. Maintenance of fixation was confirmed using the scleral search-coil technique²⁷, measuring eye position with an accuracy of 30' every 16 ms. Data were rejected from trials during which the monkey was not fixating appropriately when the stimulus appeared, or during which eye movements of more than 2° occurred in the first 600 ms following stimulus onset.

X-radiographs were used to locate the position of the microelectrode on each recording track relative to bony landmarks. The position of cells was reconstructed from the X-ray coordinates taken, together with serial 50- μ m histological sections showing the micro-lesions made at the end of some of the microelectrode tracks. Recording sites were all located within the lower bank of the superior temporal sulcus and in the adjacent dorsal part of the inferior temporal gyrus. All recording sites were localized within cytoarchitectonic areas TEa, TEM and TE3, as described previously²¹ (Fig. 7).

Stimulus presentation and task. The 32 visual stimuli (Fig. 1) were stored digitally on a computer disk, and displayed on a Sony video monitor using a Data Translation video framestore (512 \times 480 pixels; Marlboro, Massachusetts). Maximum and minimum luminances on the screen were 5.2 and 0.22 footlamberts, respectively. The exposed shape was either white on a black background, or vice versa (Fig. 1). Each shape averaged 2.8° in width and 3.5° in height, with the center of the curved contour located centrally at fixation.

Before a trial, the whole screen was gray. The screen then went black or white for 1 s, so that the subsequent figural stimulus could later appear against this background with the opposite polarity, which produces entirely unambiguous figural assignment in human observers. This preliminary change to the luminance of the whole screen was unrelated to our comparisons (see also the baseline data, before onset of the experimental stimulus, in Fig. 6). After 1 s, a central fixation dot of opposite polarity to the rest of the screen appeared for 500 ms. The fixation dot was followed by the experimental shape for 1 s, then the screen returned to gray for 3.5 s, before the start of the next trial. The monkeys performed a simple visual task during testing (adapted from ref. 29) to ensure that they fixated the stimuli. If the shape shown centrally was any one of the 32 in the experimental set (a black or white shape, Fig. 1), then the monkeys could obtain a fruit juice reward during its exposure, provided they were fixating within 1° of the central location. If the central shape was a red square (11% of trials, excluded from analyses), then the monkey had to withhold licking to avoid ingesting aversive hypertonic saline. A 0.5-s signal buzzer preceded the presentation of the stimulus. (This sounded concurrently with the central fixation point.) Thus, if the monkey fixated correctly before the stimulus appeared, he had sufficient time to discriminate black or white experimental shapes from the red square, and then obtain fruit juice while it was still available (during the central stimulus).

Before the experiment, the monkeys had been trained on a simple visual discrimination task. They viewed the monitor with a central fixation point, and could lick to obtain fruit juice when a white or black



circle was presented centrally for 1 s immediately after the buzzer, but had to withhold a lick when a red square was presented instead. Having mastered this with greater than 95% accuracy, they were trained to maintain central fixation. After several weeks of this training (without any exposure to the experimental stimuli like those in Fig. 1), the experimental trials were run in blocks of 36, each comprising the 32 experimental stimuli, plus 4 trials with red squares, all in random order. For each cell studied, three to seven blocks of trials were run, each with different random orders of stimuli.

Analyses. The two monkeys gave the same pattern of results, and so are considered jointly here. For each trial, the number of action potentials occurring in a 500-ms period starting 100 ms after stimulus onset was initially considered. This period was chosen because most of the neurons studied typically showed vigorous responses to visual stimuli with latencies just above 100 ms, and the monkeys consistently held central fixation for the first 600 ms of stimulation. To test whether a neuron was showing selectivity among the set of 32 shapes, analysis of variance was performed on the response rates to the different stimuli. Only those cells (78/88) that showed a significant effect of stimulus (at $p < 0.001$ or better) were included in the further analyses (31 from one monkey, 47 from the other), as only these could address our experimental questions.

For each of our three orthogonal transformation of the stimuli (contrast polarity, mirror-imaging and figure-ground reversal) the set of 32 stimuli can be divided into two subsets of 16, one subset providing transformed versions of each member of the other subset. To calculate the influence of one specific transformation (such as contrast reversal) on the stimulus selectivity of a single cell, we correlated the firing rates to the 16 stimuli in one subset against those for the corresponding members of the other subset (Fig. 2; data from one illustrative cell), using Spearman's rank-order correlation. This was initially done for firing rate across the 100–600-ms time bin following stimulus onset, for every cell (Fig. 5).

To see how the correlations (and thus the generalization of stimulus selectivity across a particular transform) developed over time, we next calculated the correlation coefficients for time bins of increasing extent. These were calculated for spikes in response to each stimulus in the first 20 ms, then the first 40 ms, and so on, up to 500 ms after the stimulus period. Average correlations climbed rapidly to form an asymptote around 200–300 ms after stimulus onset for contrast and mirror-image transforms, but remained near zero throughout the trial for figure-ground reversal.

As another way to study the effects of stimulus transforms on the stimulus selectivity of the cell population, we examined how the difference in responses to the preferred and the non-preferred stimulus developed over time. This was done for each cell by calculating the response to its optimal stimulus in successive 20-ms time bins. A similar time course of firing rate was then calculated for the 'non-preferred' stimulus for each cell. These values were then averaged across the population of 78 cells to produce the diagram shown in Fig. 6a. Analogous procedures were used to plot the responses to contrast-reversed versions of the same two stimuli (Fig. 6b), mirror-reversed versions (Fig. 6c) or figure-ground reversed versions (Fig. 6d).

For the human shape-judgment task, observers were presented with a sample shape for 400 ms, and two test stimuli were then added to the display at bottom left and bottom right. They were asked to judge which of these two test stimuli was more similar in shape to the sample. The test stimuli were two different shapes drawn with equal probability without replacement from the following set: the original shape, a contrast reversal, a mirror reversal, a figure-ground reversal and a shape with a different contour. Observers indicated by pressing a left or right key which test shape was more like the sample shape. The 12 repetitions of each of 20 possible permutations of test pairings were averaged together to generate overall preferences for each type of transform when presented at test.

ACKNOWLEDGEMENTS

G.C.B. was supported by grants from the National Institutes of Health (R29 NS27296) and the National Science Foundation (SBR 96-16555). J.D. was supported by the Biotechnology and Biological Sciences Research Council (UK).

RECEIVED 27 APRIL; ACCEPTED 30 JULY 2001

1. Baylis, G. C., Rolls, E. T. & Leonard, C. M. Functional subdivisions of the temporal lobe neocortex. *J. Neurosci.* 7, 330–342 (1987).
2. Desimone, R., Schein, S. J., Moran, J. & Ungerleider, L. G. Contour, color and shape analysis beyond the striate cortex. *Vision Res.* 24, 441–452 (1985).
3. DiCarlo, J. J. & Maunsell, J. H. R. Form representation in monkey inferotemporal cortex is virtually unaltered by free viewing. *Nat. Neurosci.* 3, 814–821 (2000).
4. Logothetis, N. K., Pauls, J. & Poggio, T. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5, 552–563 (1995).
5. Riesenhuber, M. & Poggio, T. *Nat. Neurosci.* 3, 1199–1204 (2000).
6. Rollenhagen, J. E. & Olson, C. R. Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science* 287, 1506–1508 (2000).
7. Rolls, E. T., Judge, S. J. & Sanghera, M. K. Activity of neurons in the inferotemporal cortex of the alert monkey. *Brain Res.* 130, 229–238 (1977).
8. Sary G., Vogel, R. & Orban, G. Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260, 995997 (1993).
9. Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* 19, 109–139 (1996).
10. Malach, R. *et al.* Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. USA* 92, 8135–8139 (1995).
11. Farah, M. J. & Aguirre, G. K. Imaging visual recognition: PET and fMRI studies of functional anatomy of human visual recognition. *Trends Cogn. Sci.* 3, 179–185 (1999).
12. Farah, M. J. *Visual Agnosia* (MIT Press, Cambridge, Massachusetts, 1990).
13. Plaut, D. C. & Farah, M. J. Visual object representation: Interpreting neurophysiological data within a computational framework. *J. Cogn. Neurosci.* 2, 320–343 (1990).
14. Rubin, E. *Visuell Wahrgenommene Figuren* (Gyldendalske Boghandel, Copenhagen, Germany, 1915).
15. Baylis, G. C. & Driver, J. One-sided edge-assignment in vision: 1. Figure-ground segmentation and attention to objects. *Curr. Dir. Psychol. Sci.* 4, 140–146 (1995).
16. Driver, J. & Baylis, G. C. One-sided edge-assignment in vision: 2. Part decomposition, shape description, and attention to objects. *Curr. Dir. Psychol. Sci.* 4, 201–206 (1995).
17. Driver, J. & Baylis, G. C. Edge-assignment and figure-ground segmentation in short-term visual matching. *Cognit. Psychol.* 31, 248–306 (1996).
18. Baylis, G. C. & Cale, E. The figure has a shape, but the ground does not: Evidence from covert testing of shape recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 633–643 (2001).
19. Hoffman, D. D. & Richards, W. A. Parts of recognition. *Cognition* 18, 65–96 (1984).
20. Baylis, G. C. & Driver, J. Obligatory edge-assignment in vision: the role of figure and part segmentation in symmetry detection. *J. Exp. Psychol. Hum. Percept. Perform.* 6, 1323–1342 (1995).
21. Selzer, B. & Pandya, D. N. Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res.* 149, 1–24 (1978).
22. Pinker, S. Visual cognition: an introduction. *Cognition* 18, 1–64 (1984).
23. Rolls, E. T. & Baylis, G. C. 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 (1986).
24. Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147 (1987).
25. Nakayama, K., Shimojo, S. & Silverman, G. H. Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18, 55–68 (1989).
26. Palmer, S. & Rock, I. Rethinking perceptual organisation: the role of uniform connectedness. *Psychon. Bull. Rev.* 1, 29–55 (1994).
27. Peterson, M. A. Object recognition processes can and do operate before figure-ground organisation. *Curr. Dir. Psychol. Sci.* 3, 105–111 (1994).
28. Robinson D. A. A method of measuring eye-movements using a scleral search coil in a magnetic field. *IEEE Trans. Biomed. Eng.* 101, 131–145 (1963).
29. Baylis, G. C., Rolls, E. T. & Leonard, C. M. Selectivity between faces in the responses of a population of neurons in the cortex of the superior temporal sulcus of the monkey. *Brain Res.* 342, 91–102 (1985).