Figure-ground mechanisms provide structure for selective attention

Fangtu T. Qiu, **Tadashi Sugihara**, and **Rüdiger von der Heydt**
Krieger Mind/Brain Institute and Department of Neuroscience, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218

Abstract

Attention depends on figure-ground organization: figures draw attention, while shapes of the ground tend to be ignored. Recent research has demonstrated mechanisms of figure-ground organization in the visual cortex, but how they relate to the attention process remains unclear. Here we show that the influences of figure-ground organization and volitional (top-down) attention converge in single neurons of area V2. While assignment of border ownership was found for attended as well as for ignored figures, attentional modulation was stronger when the attended figure was located on the neuron's preferred side of border ownership. When the border between two overlapping figures was placed in the receptive field, responses depended on the side of attention, and enhancement was generally found on the neuron's preferred side of border ownership. This correlation suggests that the neural network that creates figure-ground organization also provides the interface for the top-down selection process.

Perception tends to segregate visual images into figures and ground, and process the figure regions, but not the ground regions (Fig. 1).[1,2] Apparently, the system is able to group the visible borders at an early stage to configurations that are likely to be objects and process this information with priority. Objects can be selected by spatial filtering when they are separated ('spotlight of attention', Fig. 2a), but such a mechanism fails when objects are partially occluded by others as in everyday images. When trying to select the bottom square in Figure 2b for example, such a mechanism would select contours in the form of an L rather than a square. Observers generally perceive the square—not the L. Apparently, the visual system first subtracts the contours of the occluding object, to use only the remaining contours for further analysis, acknowledging that information is missing (Fig. 2c). Thus, an essential step in the interpretation of images is the correct assignment of the visible borders (contours, edges) to foreground regions.[3,4] Regions of occluded objects are bounded by two types of contour, those that are inherently related to the object (*intrinsic* contours), and those formed accidentally by interposition of another object (*extrinsic* contours).[3] Only the intrinsic contours should be processed together for shape recognition; extrinsic contours should be excluded. Single cell recordings from monkey visual cortex have shown that assignment of *border ownership* occurs at stages as early as areas V1 and V2.[5–7] Neurons in V2 are also influenced by top-down attention.[8–12] How these two processes are related is not clear. Is figure-ground organization the result of selective attention, or is it an independent process? If it is independent, as we shall argue, what is its role in the deployment of attention? Does it enable the attention process to select contours according to border ownership (*interface hypothesis*), or is attentional modulation determined merely by the distance of a contour from the focus of attention (*spatial attention hypothesis*)?

Correspondence: Rüdiger von der Heydt, Krieger Mind/Brain Institute, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, Phone: 410 516-6416, Fax: 410 516-8648, E-mail: von.der.heydt@jhu.edu.

Results

We studied the responses of neurons in area V2 under conditions when monkeys performed a shape discrimination task that required selective attention (Fig. 2d–f). At the beginning of a trial, the animal was required to fixate a cross in the center of a display. After a short delay, three figures were simultaneously displayed that could be squares or trapezoids, and the animal had to focus on one of the figures (the target) and report its shape with an eye movement (monkey TE) or hand movement (monkey LA; see Methods). Each of the three figures could be a square or a trapezoid with 50% probability; however, the animal was rewarded only if it responded correctly to the target figure. The target was specified by instruction displays at the beginning of each block of trials (Fig. 2f). Which of the three figures was the target varied between blocks (typically 20 trials). For example, in one block the target could be the middle figure (Fig. 2f, left), and in the next block it could be the right-hand figure (Fig. 2f, right). In the subsequent test trials there were no cues that would differentiate the target from the other figures and the monkeys had to remember which figure should be attended.

In one set of experiments, the three figures were separated (Fig. 2d), in another set, two of the figures overlapped (Fig. 2e). An edge of one of the figures was placed in the receptive field (RF) of the neuron under study. In the case of separated figures, border ownership was varied by flipping each figure about one of its edges, and at the same time interchanging the colors of figures and background (compare top and bottom displays in Fig. 2d). Displays with reversed contrast were also tested (not illustrated). In the case of overlapping figures, border ownership was varied by changing the order of occlusion (Fig. 2e). Note that in each case, the local stimulus in the RF was the same for the two border ownership conditions. Border ownership was varied randomly between blocks or from trial to trial.

1. Border ownership assignment in the absence of attention

The purpose of the first experiment (Fig. 2d) was to see if border ownership selectivity of a neuron was affected depending on whether attention was directed to the figure stimulating the neuron or to another figure in the display. Significant modulation by the monkey's attention was found in 80 of 243 cells as a main effect ($p < 0.05$, analysis of variance—ANOVA; see Methods). In 64 cells attention enhanced the responses, in 16 it reduced them. Sixty-two cells showed main effects of both attention and border ownership, and 60 cells showed significant interaction between the two factors (with or without main effects). Taking together all cells that showed both main effects or interaction, we counted 100/243 neurons (41%) in which attention and border ownership influences converged (Fig. 3a, cross hatched). In contrast, 91 cells (37%) showed significant border ownership modulation without any influence of attention (Fig. 3a, striped vertically). The existence of these cells indicates that attention is not required for border ownership modulation.

How does attention affect border ownership modulation in the cells in which both influences converge (cross hatched sector in Fig. 3a)? We compared the mean response strengths in the four experimental conditions in this group of cells (Fig. 3b). We illustrate the two border ownership conditions by depicting the receptive fields of two neurons with opposite side preference (ellipses with arrows) in one of the two stimulus configurations (this is justified because the distribution of border ownership selectivity is isotropic, see Supplementary Fig. S1). The border-ownership related response difference was nearly as strong for the ignored figure as for the attended figure (modulation index 0.22 versus 0.25). Also when averaged over all cells tested, border-ownership modulation was similar for the two attention conditions (0.10 and 0.11). There was no difference in latency between the border ownership signals for attended and ignored figures (Supplementary Fig. S2), which argues against the possibility that the signals might be generated by a serial attention process.

We conclude from this experiment that border ownership signals can emerge without the influence of attention and that the overall strength of border ownership modulation is nearly the same for figures that the monkey tries to ignore as for figures at the focus of its attention.

2. Attention and extrinsic border suppression

Next we examined configurations in which two figures overlapped (Fig. 2e). The border between the two figure regions (the occluding edge) was placed in the RF of the neuron under study and responses were recorded for the two directions of occlusion and for attention on foreground figure and attention on background figure. If border ownership assignment provides the structure for selective attention, then the occluding edge should not be selected when attention is on a background figure (*interface hypothesis*). The *spatial attention hypothesis*, if anything, predicts the opposite, because the occluding edge is closer to the focus of attention in the attend-back condition (see pictograms in Fig. 3e); the saccades after the fixation period indicated that the focus of attention was approximately in the center of the attended figure (Supplementary Fig. S3).

We found that varying attention had a strong effect: Of 216 cells tested, 103 (48%) showed an influence of the side of attention, and 66 of these (31% of the total) showed both influences combined, in form of significant main effects or interaction (Fig. 3c). In the population response, border ownership modulation was strong when attention was in front, but close to zero when attention was in back (Fig. 3d). The responses to the occluding edge were lower in the attend-back condition than in the attend-front condition, in support of the interface hypothesis (Fig. 3e). We quantified the relative attenuation of the extrinsic border by calculating, for each neuron, the ratio of the mean firing rate for the attend-back condition to the mean firing rate for the attend-front condition, as symbolized by the division operator in Figure 3e. The distribution of this ratio for the border ownership neurons corresponding to the attended region (Fig. 3e, black symbol) shows attenuation in nearly all cases (61/66, $p < 10^{-10}$, proportion compared to 1/2, large sample test). The median response ratio was 0.72, and the most selective cell showed a ratio of 0.22. Thus, border ownership mechanisms attenuate extrinsic edge signals in selective attention. The attenuation may not appear as strong as expected from perception. Perhaps further suppression is achieved by similar mechanisms at subsequent stages, for example in V4, where border ownership is also represented.[5] How the degree of edge suppression is to be calculated depends of course on the way the V2 signals are 'read' by a subsequent form recognition process. It is conceivable that this process reads not the responses, but the border ownership signals, i.e., the response *differences* between cells with opposite border ownership preference (black and gray symbols in Fig. 3e). Like the responses, the border ownership signal also carries orientation and color information.[5,13] Calculated this way, suppression of the extrinsic edge signal would be total (median signal ratio −0.05).

A comparison of Figure 3b and Figure 3d shows that the net effect of switching attention was weaker for separated figures than for overlapping figures, whereas the opposite was true for border ownership modulation. This dissociation is in agreement with previous studies showing border ownership modulation to be stronger for contours of isolated figures than for contours between overlapping figures,[5] and attentional modulation to be weaker for objects that are widely separated than for objects that are 'competing' within the same receptive field.[9]

3. Time course of border ownership and attention effects

Border ownership and attention effects both emerged with only a short delay after the beginning of stimulus evoked activity in V2 (Fig. 4). We plotted the average differences in firing rate caused by varying border ownership (blue dashed line) and site of attention (red solid line) for the displays with separated (middle) and overlapping figures (bottom). The border ownership

signal for single-figure displays is also shown for comparison (top). For separated figures, the attention effect shows the difference between attending to a figure at the RF and attending to a distant figure. For overlapping figures, it shows the response difference produced by changing side of attention relative to the RF, which is discussed in the next section. For comparison we also plotted the time course of the mean responses (black, right scale). Border ownership as well as attention differences begin to emerge around 50 ms after stimulus onset. In the case of separated figures, the attention modulation increases gradually up to about 180ms, while the border ownership signal has a steep onset followed by a phase of relative constancy or decline (similar to the curves for single figure). In the case of overlapping figures, the attention modulation is stronger and has a steeper initial slope. The early onset of the difference signals in Figure 4 shows that the processing is surprisingly fast, considering that border ownership assignment involves integration of image context, and that attentional modulation in addition depends on a central process. Because redirecting attention in response to a visual cue takes 150ms or more[14] the attention effects observed in Figure 4 must be 'programmed' already before the onset of the stimuli (the target figure was specified at the beginning of each block of trials). The fast onset of attentional modulation will be discussed below in the context of a model.

4. Common circuits for figure-ground and attention

The test with overlapping figures revealed an asymmetry of the attention effect which can be seen in typical example neurons in Figure 5. Border ownership and side-of-attention were varied independently (see insets). For the illustration we assumed that the RF orientations were vertical and border ownership preference was left. Thus, left figure in back is the non-preferred condition (**a–b**, dashed curves) and left figure in front is the preferred condition (**c–d**, solid curves). Asterisks and color indicate side of attention. The curves at the bottom show the mean firing rates for the example neurons (**e**). It can be seen that attention on the left figure (red) enhanced responses compared to attention on the right figure (blue), irrespective of the occlusion condition (red curves tend to be higher than corresponding blue curves, see also population averages in **f**). Thus, in both examples, the attention effect was asymmetrical about the RF, and the side of attention enhancement was the same as the preferred border ownership side. The similar ordering of the population curves (**f**) indicates that this is a consistent pattern. Since border ownership and side of attention are independent factors, there is no a priori reason why side of attention enhancement should be correlated with preferred side of border ownership.

The population results show that there was a correlation (Fig. 6). For each neuron we determined the main effects of the two factors, as given by ANOVA, and plotted the degree of border ownership modulation (vertical axis), and the response modulation produced by side of attention (horizontal axis). A positive value of side-of-attention modulation means that the side of enhancement was the same as the preferred border ownership side, a negative index means that it was opposite. Of the neurons that combined influences of border ownership and attention (filled circles, histogram at top) a majority (72%) showed attentional enhancement and border ownership preference for the same side (symbols in right half). Moreover, neurons with large border ownership modulation tended to show also large side-of-attention modulation. The mean side-of-attention index was 0.128, which was significantly different from zero ($p = 2.4 \cdot 10^{-5}$, N = 66, t-test). There was a significant shift to positive values also in the entire population (mean = 0.047, $p = 1.6 \cdot 10^{-4}$, N = 215). Thus, the test with overlapping figures reveals that attentional modulation is spatially asymmetric about the RF, and this asymmetry is correlated with border ownership preference. [In some neurons, responses to the occluding edge were enhanced by attention on the foreground compared to attention on the background irrespective of border ownership. However, front-back attention modulation was overall weaker than side-of-attention modulation (Supplementary Fig. S4).]

After observing spatial asymmetry of attentional modulation in this experiment we re-examined the data of experiment 1 (separated figures) and found a similar asymmetry: the attentional enhancement was stronger for figures on the preferred border ownership side than figures on the nonpreferred side (modulation index 9.8% vs. 3.0%, N = 100, t = 4.11, P = 0.0001, paired t-test). Thus, even in the case of separated figures, when spatial mechanisms would be adequate for attentional selection, the neurons are more susceptible to attentional modulation on the side of border ownership preference than on the other side. These observations have implications for the mechanisms underlying selective attention that will be discussed below.

5. Controls and comments

The results just described were obtained in two animals performing somewhat different tasks (see Methods): TE signaled the shape of the target figure by making different saccades, whereas LA responded manually. We used a different task in the second animal (LA) to make sure that the attention effects we had observed in TE were not a result of training the animal to make specific eye movements. The results from the two monkeys were virtually the same in every respect: strength of attention effect and similarity of border ownership modulation with and without attention in the separated figure condition; spatial asymmetry of attention modulation (Fig. 4–Fig. 5) and correlation between side-of-attention and border ownership modulation in the overlapping figure condition (Fig. 6). Also the distributions of saccades to the target figure were quite similar, despite the fact that LA was not required to make any specific eye movements (Supplementary Fig. S3). This shows that LA processed the stimulus in the same way as TE did, and that training an animal to respond by eye movements did not alter the way attention modulates the visual responses in V2.

Regarding the experiment with overlapping figures, it might be argued that the monkey focused attention on the location of the hidden edge when the background figure was the target, in which case its focus of attention would not have been exactly on the RF because this was centered on the visible, occluding edge. Two observations show that this cannot be the explanation for the asymmetry of the attention effect. First, the distribution of post-fixation saccades indicates that attention was directed to the center of the target figure and not to the hidden edge (Supplementary Fig. S3). Second, attention modulation showed the same spatial asymmetry irrespective of the direction of occlusion (Fig. 5).

To address the question of whether variations in fixation behavior could have contributed to the correlations shown in Figure 6, we analyzed the position of gaze during the 200 ms window used for the analysis of neuronal responses. Eye movements could mimic attention effects only if the fixation position would differ systematically between attention conditions, for example so that the test edge would be centered in the RF in one condition, but displaced in the other. We found no such differences, neither for individual neurons, nor in the population means (Supplementary Fig. S5). Thus, we can rule out eye movements as a source for the correlations seen in Figure 6.

We also ruled out the distribution of the RF orientations and positions of the neurons in our sample as a source of correlation. Attention modulation did not depend on the orientation of the stimulus axis (which varied with RF orientation), or the orientation of the stimulus axis relative to the vector connecting RF and fovea (Supplementary Fig. S1). This means that attention modulation was not related to the location of the figures relative to the center of fixation.

We further considered the possibility that aspects of the task other than attention could have influenced the results. Because we used squares and trapezoids, the orientation of the edge in the RF varied slightly (typically ±7deg). However, the different shapes of figures contributed

equally to the responses in each experimental condition, and there was no interaction between shape and site of attention (14 of 253 cells showed interaction at $p < 0.05$, not different from proportion expected by chance, $p = 0.71$).

6. Understanding the mechanisms

The asymmetry of receptive fields regarding attentional modulation was an unexpected finding. A plausible explanation for this asymmetry is that the same circuits that produce border ownership modulation also provide a structure for attentional selection. We have previously proposed a model for border ownership assignment based on simple circuits that integrate image context.[15,16] If we assume that top-down attention works by activating the same circuits, then all the above findings fall into place.

The principle of this idea is illustrated in Figure 7a–b, which shows stimuli and receptive fields at the top, and the corresponding cortical neurons below. The black dots represent border ownership selective V2 neurons as recorded here ('B cells'), the larger dots corresponding to the neurons whose receptive fields were stimulated by edges of the figures (ellipses with arrows). Opponent pairs of B cells are encircled (red dashed lines). We assume that border ownership selectivity is created by grouping cells ('G cells', numbered hexagons) which connect B cells in a roughly co-circular arrangement of receptive fields. The G cells sum the signals of the B cells and, via feedback connections, increase the gain of the same B cells (for graphical clarity only one line is shown for the two directions of connectivity). Each B cell is connected asymmetrically to G cells on one side and is therefore facilitated only when a figure activates a grouping cell on that side.

The grouping cell network is the key to understanding the interplay between attention and figure-ground organization. We assume that selective attention excites G-cells at the focus of attention, or inhibits G-cells surrounding it (Fig. 7c–d). The spatial asymmetry of the attention effect and its correlation with border ownership are then obvious corollaries: because border ownership preference of a neuron is determined by the same connectivity, side of attentional enhancement and preferred border ownership side must be the same; the responses of a B-cell are enhanced if the focus of attention is on the side of its G-cell connection. For example, in the case of overlapping figures, attending to the right figure means injecting activity into G-cell 1 (Fig. 7d top), which will enhance the responses of the connected B-cell (black RF). The fast onset of attentional modulation in the overlap experiment (Fig. 4 bottom, red solid line) is explained naturally by the model: because feedback from G-cells sets the gain of B-cells, the differential activation of the G-cells on either side sets different gains in the opponent B-cells, producing a firing rate difference right from the beginning of the response to the stimulus. That G-cells set the gain of B-cells, but do not excite or inhibit them, is apparent in Figure 4 which shows that varying attention did not produce differences in baseline firing rate (initial segments of curves are close to zero).

Also the details of the results shown in Figure 3 follow from the scheme of Figure 7. In the case of overlapping figures, when attention is on the front figure (Fig. 7d top), border ownership modulation is strong (cf. Fig. 3d), because the positive effects of configuration and attention cumulate in G-1, boosting the responses of the right-pointing B-cell (black), but when attention is on the background figure (Fig. 7d bottom), G-2 is activated by attention, whereas G-1 is favored by configuration, and both effects cancel in the difference (*extrinsic edge suppression*, cf. Fig. 3d). In the case of separated figures, attention on G-2 (Fig. 7c bottom) will enhance the activity of the left-pointing B-cell (black RF) while attention on G-3 will reduce it (Fig. 7c top). No effect is predicted on the partner B-cell (gray RF) because its grouping cell G-1 is in the shadow of attention in both conditions. As mentioned at the end of section 4, the attentional enhancement was in fact 3-times weaker on average when the figure was on the non-preferred than the preferred border ownership side. (That it was not zero can

be explained plausibly by assuming that the attention effect is a combination of the grouping cell mechanism, as proposed, and a spotlight mechanism.)

The model accounts for three aspects of the results described above. It explains how the system uses image context to generate border ownership signals; it explains the spatial asymmetry of the attention influence; and it explains why the side of attention enhancement is generally the same as the preferred side of border ownership. The existence of G cells is hypothetical as yet. Our results suggest that border ownership preference is a fixed property of the neurons, implying that G-cells are pre-established (by genetic or experiential factors). Our model postulates that G cell templates come in a range of sizes and cover the visual field densely, but with relatively coarse spatial resolution. Their resolution should be comparable to that of attention,[17] which is much lower than the visual acuity and the resolution of the receptive fields of V1. This means that the model can function with a relatively small number of G-cells, about 1% of the cells representing image information.[16] G-cells might reside within or outside V2. As pointed out,[5,16] the short latency of the border ownership signal (Fig. 4) and the slow conduction of intra-cortical horizontal fibers argue for a location in a higher-level area such as V4, because the propagation of image context information can then be achieved via myelinated fibers which are an order of magnitude faster than intracortical fibers.[18] V4 appears as a likely candidate also because it receives direct projections from the parietal area LIP which is an important stage in the generation of top-down visual attention.[19]

Discussion

Our results demonstrate that selective attention and figure-ground organization involve overlapping populations of neurons in V2. A fraction of the cells showed border ownership selectivity without any attention modulation, others exhibited an influence of attention without border ownership selectivity, and a large fraction (about 40% of the cells) showed both influences combined. Such a result would be expected if border ownership assignment and attentional modulation were produced by two independent mechanisms that interact at this stage.

We show that border ownership assignment occurs simultaneously for multiple figures in the display, including figures outside the focus of attention (Fig. 3a). Figure-ground coding has also been observed in the form of enhancement of responses to the inside of a figure relative to the outside,[20–22] which is also independent of attention.[12] The question of whether perceptual organization occurs at preattentive levels has been debated since the early Gestalt writers.[1,2,23–25] Studying preattentive processing in psychological experiments poses a conundrum because instructing subjects to make a judgment about some aspect of a stimulus display seems to require attentive processing, and judgments about non-attended aspects would have to rely on memory. However, there could be preattentive processing that does not leave a trace in memory. Our results show that border ownership was assigned for any figure in the display, whether it was attended or not. The possibility that the figures were elaborated sequentially by a fast serial attention process is unlikely, because there was no latency difference between the border ownership signals for attended and ignored figures. Thus, as far as human brain processes can be inferred from a study of monkey brains, we can say that figure-ground organization, as known from human perception, does occur preattentively.

The convergence and largely additive effects of figure-ground and attention effects in single neurons of the V2 edge representation also explain the observation that perceived figure ground organization can be influenced by attention[1] (inspection of Fig. 1 shows that the shape of a region that is rendered background by bottom-up mechanisms, and therefore not recognized immediately, can be clearly perceived with volitional attention). In the model of Figure 7, G-

cells can be activated by a configuration of edges, or by top-down attention, and both modes of activation are equivalent in raising the gain of B-cells in the edge representation.

The most telling result of our experiments is the asymmetry of V2 receptive fields with respect to attentional modulation and its correlation with the border ownership preference of the cells. This correlation indicates that top-down attention processes share neural circuitry with the mechanism underlying context integration in figure-ground organization. Asymmetry with respect to attentional modulation has been demonstrated in receptive fields of V4 neurons,[26] a finding that might be related to the mechanism we describe here.

The attention effects in our experiments might be interpreted as examples of 'biased competition' in which visual objects compete for neural representation and top-down attention can bias the competition in favor of one or the other object.[27,28] The attentional modulation of the neural responses to the border between figures (Fig. 4–Fig. 5) might reflect a competition between two objects, similar to the one that occurs when two separate bars are presented within a receptive field.[9] Critical issues in the biased competition theory involve the questions of how the system determines what is an object, how it forms object representations, and how top-down attention signals can address these representations. The present results point to specific circuits in the visual cortex that bind features to larger compounds and provide a structure for selective attention.

Methods

We studied two adult rhesus monkeys (Macaca mulatta), one male and one female. The details of our general methods have been described.[5,6] The animals were prepared by implanting, under general anesthesia, first three small posts for head fixation, and later two recording chambers (one over each hemisphere). Behavioral training was achieved by controlling fluid intake and using small amounts of juice or water to reward correct responses. All animal procedures conformed to US National Institutes of Health and USDA guidelines as verified by the Animal Care and Use Committee of the Johns Hopkins University.

Recording

Single-neuron activity was recorded extracellularly with epoxy-insulated tungsten microelectrodes inserted through the dura mater within small (3–5 mm) trephinations. Area V2 was identified by its retinotopic organization and by histological reconstruction of the recording sites, as described.[5] Action potentials were discriminated using a spike sorting device (Alpha Omega). Only isolated single unit activity was analyzed. Receptive fields were in the lower hemifield at eccentricities ranging between 0.75 and 12 deg (median 2.2 deg). Eye movements were recorded for one eye using an infra-red video based system (Iscan ETL-200) with a resolution of 5120 (H) and 2560 (V). The eye was imaged through a hot mirror (a mirror that selectively reflects infrared), with the camera placed on the axis of fixation. The optical magnification in our system resulted in a resolution of the pupil position signal of 0.03 deg visual angle in the horizontal and 0.06 deg in the vertical. However, noise and drifts of the signal reduced its accuracy.

Behavioral tasks

Animals performed two tasks, a shape discrimination with initial fixation, and a simple fixation task. Shape discrimination was taught first. Upon appearance of a fixation spot, the animal could initiate a trial by fixating the spot, which was detected by monitoring the eye movements. After fixation was maintained for 0.3 s, a figure was displayed that could be a square or a trapezoid, and the animal was rewarded if it signaled the shape correctly. Monkey TE responded by making a saccade to the figure if it was a trapezoid, and looking off the screen if it was a

square. A trial was rewarded only if the first saccade after fixation landed in the correct target zone.[29] Monkey LA responded by pulling or pushing a lever. A correct response terminated the trial. After an incorrect response, the trial was terminated and a 3 s delay ensued. Upon termination of a trial, the screen was blanked for 1–1.5 s (plus the additional delay after an error) until the fixation spot came on again and a new trial was enabled. Once the animals performed the shape discrimination reliably, two additional figures were added and the animals were trained to perform the task with one of the figures, the target, as specified by instruction trials at the beginning of each block of trials. In these trials the target figure was shown as solid and the other figures as outlines. Which of the figures was the target varied between blocks. The shape of each figure varied randomly from trial to trial. Once the animals mastered the task with three spatially separated figures, a variant of the display was introduced in which two of the figures partially overlapped. The blocking of trials and the sequence of events in each trial were the same, except that a certain time after stimulus onset the top (occluding) figure was moved so as to expose the bottom figure completely. This occurred after 0.5 s for monkey TE and after 0.2 s for monkey LA. Thus, in trials in which the bottom figure was target, correct performance required that the animal waited until that figure was exposed before responding. Both monkeys performed the tasks well above chance level (TE, 80%, LA, 91% correct). To check if the responses of monkey LA were based on processing the stimulus during the fixation period (in TE this was obviously the case, because he responded with a saccade at the end of the fixation period), we modified the display sequence in some of the training sessions so that the display was blanked when a saccade was detected. In these sessions, in which post-saccadic information could not be used, LA's performance was also well above chance (72%).

The animals learned also to perform a fixation task in which trials were rewarded only if the eye position signal stayed within a window of 0.75 deg radius for 2s. (Note that this scheme actually produces more accurate fixation than suggested by the size of the reward window because noise and drifts of the eye movement signal effectively produce a negative gradient of reward probability away from the fixation point.) This task was used for mapping of receptive fields, for the general characterization of selectivity, and for the standard test of border ownership. The shape of the fixation spot told the animal which task to perform.

Design and presentation of stimuli

Stimuli were generated on a Pentium 4 Linux workstation with NVIDIA GeForce 6800 graphics card using the anti-aliasing feature of the Open Inventor software, and were presented on a 21-inch EIZO FlexScan T965 color monitor with 1600×1200 resolution, a 100 Hz refresh rate, and a maximum luminance 93 cd/m^2. Background luminance was 28 cd/m^2 except for conditions in border ownership tests in which figure and background color were flipped. The display was viewed binocularly at a distance of 100 cm and subtended 22.7 by 17.1 deg of visual angle. Stationary bars were used to determine the color preference, and bars and drifting gratings to map the minimum response field of each cell. Orientation tuning curves were recorded using moving bars.

Three shapes of figures were used in the main experiments, a square, and two trapezoids that were derived from the square by tilting one side (*A*) either clockwise or counterclockwise, typically an angle of 7 deg. The figures typically measured 3 deg on a side, but smaller figures were often used in foveal cells. All figures had rounded corners (radius 9% of figure size) to avoid the use of angles as a cue in the task. For the overlapping figures, the amount of overlap was about 13%, and the figures were displaced parallel to the occluding edge by about 9% of the figure size. In each trial, three figures were simultaneously presented with the shape of each figure chosen randomly to be a square with probability 0.5, or either kind of trapezoid with probabilities 0.25. The figures were presented with orientation of side *A* for the square shape equal to the preferred orientation of the cell under study. The centers of the sides *A* were

arranged on a circle around the fixation point. The spacing depended on the size of the figures and was typically 60 deg polar angle. In experiments with separated figures, border ownership was varied by flipping each figure about side *A* (see Fig. 2d). This variation was also block-randomized (trial-by-trial randomization was not used because changing the positions made it difficult for the monkeys to perform the task, since they remembered the target figure by its location). In experiments with overlapping figures, border ownership was varied by switching the direction of occlusion (which figure was in front and which in back, Fig. 2e). In about half of these experiments, border ownership was randomized trial by trial. In this variant, the central (occluding or occluded) edges were tilted to make trapezoids. Each configuration could be presented with two contrast polarities so that the edge in the RF assumed either polarity of contrast (e.g., light–dark and dark–light). Both polarities were tested in all of the separated-figure experiments, and in most of the overlapping-figure experiments. Figure contrast was randomized trial by trial. In both experiments, a total of 5 factors were varied factorially: site of attention, border ownership, local contrast, shape, and direction of tilt (for trapezoids).

Procedure

Upon isolating a cell we first characterized its selectivity for color, bar size, and orientation, and mapped its RF.[5] A standard test of border ownership with a single square, using square sizes of 3 and 8 deg,[6] was also performed in most cells. The fixation paradigm was used for these basic tests. Subsequently, one of the selective attention experiments (or both, if time permitted) was performed using the shape discrimination paradigm. Each of the two attention and two border ownership conditions was typically presented 40 times, one per trial. Our sample is not biased with respect to the effect of attention. However, because neurons were usually selected for the main tests after the standard border ownership test was performed, the proportion of border-ownership selective cells in our sample (74%) was higher than average. Among the total of 666 cells in which the standard test was performed, 303 (45%) were found to be border ownership selective. This is virtually the same as the proportion of 184/423 (43%) found with the same test in experiments in which the animal was never trained to pay attention to the stimuli, but, on the contrary, its attention was engaged at the fovea by a demanding fixation task (stereoscopic adjustment within a small fixation target).[5–7] This is important, because it shows that the overall frequency of border ownership selective cells was not altered by training the attention task in the present study.

Data analysis

The spike activity during periods of 200 ms after stimulus onset was analyzed. We chose this interval because eye movement recordings indicated that no systematic shifts of gaze occurred before 160 ms (Supplementary Fig. S5). Because V2 neurons respond with a delay of about 40ms,[5] we can assume that eye movements that occurred after 160 ms did not influence the activity during the analysis period. Neurons that responded with less than 4 spikes/s mean firing rate in each of the four border ownership/attention conditions were excluded because we felt that our stimuli were not appropriate for these cells (10%). Analysis of variance (ANOVA) was performed on the square-root transformed spike counts. This transformation serves to homogenize the variances and produces approximately normal distributions. The ANOVA included five factors: site of attention, border ownership, local contrast, shape, and direction of tilt (nested within shape), and was performed on each neuron. The main effects of attention and border ownership and their interaction are discussed in this paper. In Figure 6, the effects were expressed by the modulation index, $M = (a-b)/(a+b)$, where *a* and *b* are the mean firing rates for the two levels of a factor. *M* is bounded within (−1, +1). For border ownership, *a* represents the preferred side, so that in this case $M \geq 0$. The population means and their S.E. for the four conditions represented in Fig. 3 were estimated by repeated measures ANOVA performed on the data from all neurons for each of the two experiments. To plot the time courses of border ownership and attention effects (Fig. 4), the differences between post-stimulus time

histograms (1 ms bin width) were calculated for all neurons that showed the influence of both attention and border ownership (main effects or interaction), subtracting the non-preferred from the preferred condition for each neuron. The difference histograms of the individual neurons were weighted by their inverse S.D. (square root of the mean squared error of the ANOVA described above) and averaged, and smoothed with a Gaussian kernel of $\sigma = 5$ms. The curves for separated figures and overlapping figures are based on different samples of cells, with some cells included in both (total number of cells, 96 for TE, 56 for LA), and the curves for single figure are based on those cells of the combined sample for which data for the corresponding size of square are available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Rubin, E. Visuell wahrgenommene Figuren. Copenhagen: Gyldendals; 1921.

2. Koffka, K. Principles of Gestalt Psychology. New York: Harcourt, Brace and World; 1935.

3. Nakayama K, Shimojo S, Silverman GH. Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. Perception 1989;18:55–68. [PubMed: 2771595]

4. Driver J, Baylis GC. Edge-assignment and figure-ground segmentation in short-term visual matching. Cogn. Psychol 1996;31:248–306. [PubMed: 8975684]

5. Zhou H, Friedman HS, von der Heydt R. Coding of border ownership in monkey visual cortex. J. Neurosci 2000;20:6594–6611. [PubMed: 10964965]

6. Qiu FT, von der Heydt R. Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. Neuron 2005;47:155–166. [PubMed: 15996555]

7. Qiu FT, von der Heydt R. Neural representation of transparent overlay. Nat. Neurosci 2007;10:283–284. [PubMed: 17310247]

8. Motter BC. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J. Neurophysiol 1993;70:909–919. [PubMed: 8229178]

9. Reynolds JH, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4. J. Neurosci 1999;19:1736–1753. [PubMed: 10024360]

10. Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J. Neurophysiol 1997;77:24–42. [PubMed: 9120566]

11. Bender DB, Youakim M. Effect of attentive fixation in macaque thalamus and cortex. J. Neurophysiol 2001;85:219–234. [PubMed: 11152722]

12. Marcus DS, Van Essen DC. Scene segmentation and attention in primate cortical areas V1 and V2. J. Neurophysiol 2002;88:2648–2658. [PubMed: 12424300]

13. Friedman HS, Zhou H, von der Heydt R. The coding of uniform color figures in monkey visual cortex. J. Physiol. (Lond) 2003;548:593–613. [PubMed: 12611925]

14. Motter BC. Neural correlates of feature selective memory and pop-out in extrastriate area V4. J. Neurosci 1994;14:2190–2199. [PubMed: 8158265]

15. Schuetze H, Niebur E, von der Heydt R. Modeling cortical mechanisms of border ownership coding. J. Vision 2003;3/9:114.

16. Craft E, Schuetze H, Niebur E, von der Heydt R. A neural model of figure-ground organization. J. Neurophysiol 2007;97:4310–4326. [PubMed: 17442769]

17. Intriligator J, Cavanagh P. The spatial resolution of visual attention. Cognit. Psychol 2001;43:171–216. [PubMed: 11689021]

18. Girard P, Hupe JM, Bullier J. Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. J. Neurophysiol 2001;85:1328–1331. [PubMed: 11248002]

19. Goldberg ME, Bisley JW, Powell KD, Gottlieb J. Chapter 10 Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior. Prog. Brain Res 2006;155PB:157–175. [PubMed: 17027387]

20. Lammea VAF. The neurophysiology of figure-ground segregation in primary visual cortex. J. Neurosci 1995;15:1605–1615. [PubMed: 7869121]

21. Zipser K, Lamme VAF, Schiller PH. Contextual modulation in primary visual cortex. J. Neurosci 1996;16:7376–7389. [PubMed: 8929444]

22. Lee TS, Mumford D, Romero R, Lamme VAF. The role of the primary visual cortex in higher level vision. Vision Res 1998;38:2429–2454. [PubMed: 9798008]

23. Mack A, Tang B, Tuma R, Kahn S, Rock I. Perceptual organization and attention. Cognit. Psychol 1992;24:475–501. [PubMed: 1473332]

24. Moore CM, Egeth H. Perception without attention: evidence of grouping under conditions of inattention. J. Exp. Psychol. Hum. Percept. Perform 1997;23:339–352. [PubMed: 9103998]

25. Driver J, Davis G, Russell C, Turatto M, Freeman E. Segmentation, attention and phenomenal visual objects. Cognition 2001;80:61–95. [PubMed: 11245840]

26. Connor CE, Preddie DC, Gallant JL, Vanessen DC. Spatial attention effects in macaque area V4. J. Neurosci 1997;17:3201–3214. [PubMed: 9096154]

27. Grossberg S. How Does A Brain Build A Cognitive Code. Psychol. Rev 1980;87:1–51. [PubMed: 7375607]

28. Desimone R, Duncan J. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci 1995;18:193–222. [PubMed: 7605061]

29. Schiller PH. The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. Vis. Neurosci 1993;10:717–746. [PubMed: 8338809]

Figure 1.
Perception tends to segregate the optical image into figure and ground regions. Figure regions seem to attract attention, and their shapes are easily recognized (e.g., the letter **F**), whereas ground regions are generally ignored and their shapes are not recognized (the light-colored letter **G** to the left of the **F**). The concept of 'border ownership' is useful in understanding this peculiar way of perception. The **G** shaped region does not own its borders: the outer borders are assigned to the gray frame and the letter **F**, the inner borders to the red **C** shape. In contrast, the **F**-shaped region owns its border completely. Single cell recordings have shown that border ownership is represented early on in the visual cortex.
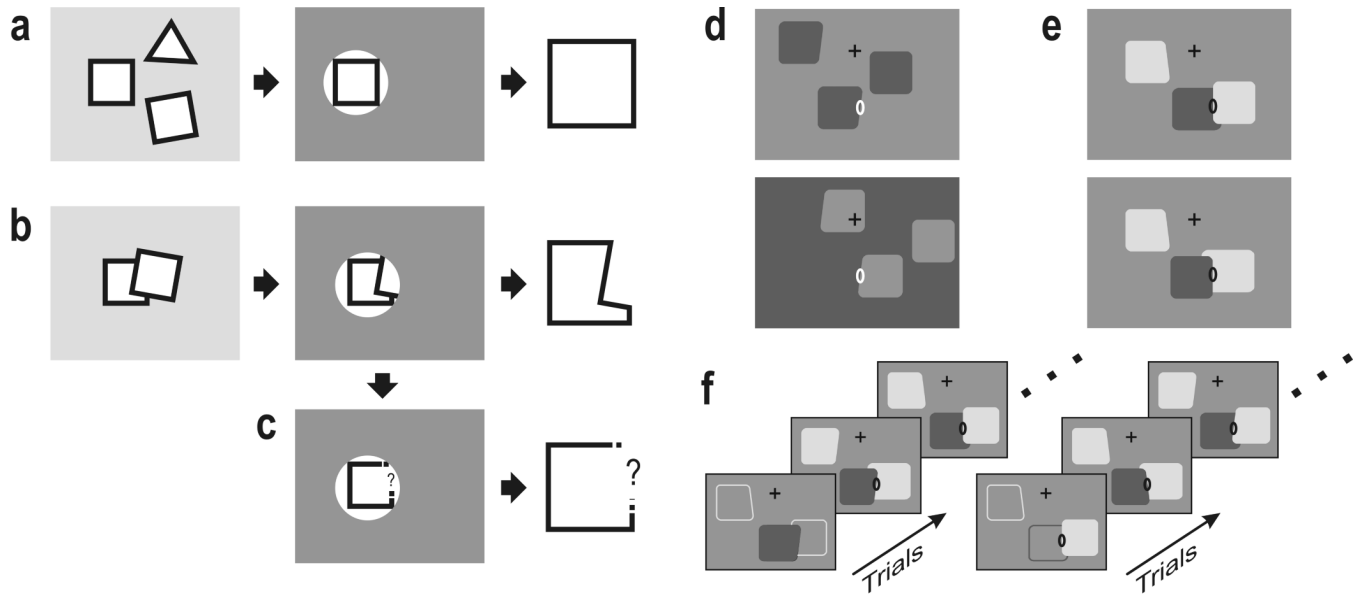
Figure 2.
Motivation and experimental design. (**a–c**) The problem of understanding images of cluttered scenes. It is easy to process one object out of several when objects occupy separate regions: a spatial selection mechanisms is sufficient (**a**), but when objects overlap, a spatial selection mechanism may not work: when trying to single out a partially occluded object, such a mechanism extracts the wrong shape (**b**). The system first needs to assign border ownership and remove the *extrinsic borders* (borders produced by occluding objects) before passing on the remaining information to a subsequent recognition stage (**c**). (**d–f**) Stimuli and behavioral task used in the present experiments. In each trial, three figures were simultaneously presented, one of which stimulated the neuron under study (ellipse, receptive field (RF); cross, fixation point). Each figure could be a square or a trapezoid. (**d**) Experiment 1, spatially separated figures. Border ownership was varied by placing a figure left and right of the RF (top and bottom displays). Note same contrast of edge in RF. (**e**) Experiment 2, overlapping figures. The occluding edge was centered in the RF and order of occlusion was varied. (**f**) Control of attention. For each presentation of figures, the monkey had to signal the shape (square or trapezoid) of one of the figures, as designated at the beginning of each block of trials by instruction displays (bottom). In the sequence on the left, the middle figure was the target for the task; in the sequence on the right, the right-hand figure.
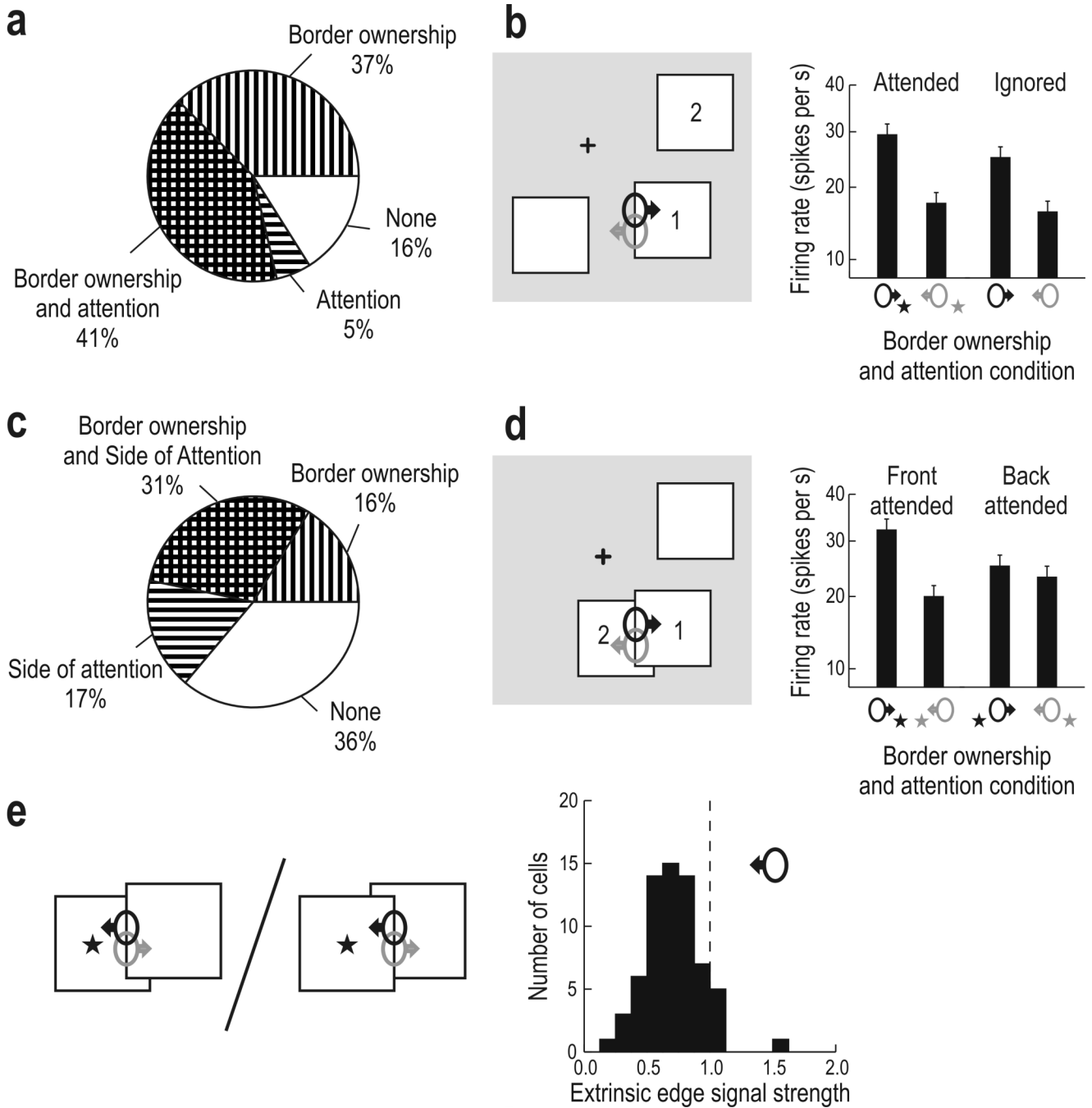
Figure 3.
Convergence of border ownership and attention influences in V2 neurons. (**a**) Frequency in percent of finding significant (p<0.05) main effects or interaction of the two factors in experiment 1 (N = 243). The cross-hatched sector represents cells in which both influences were combined (both main effects or interaction). (**b**) Mean response to the left edge of figure **1** in neurons coding right border ownership (**black**) and left border ownership (**gray**) when attention was focused on figure **1** (**attended**) and on figure **2** (**ignored**). Border ownership assignment occurs for attended as well as ignored figures. (**c–d**) Border ownership coding for overlapping figures when foreground figure (**1**) or background figure (**2**) is attended (N = 216). Border ownership modulation is strong when attention is on foreground, but weak when

attention is on background. Bars represent means, error bars s.e.m., of square-root transformed spike counts, averaged over all cells that showed significant influences of attention and border ownership (in form of main effects or interaction). Nonlinear scale reflects square-root transform (see Methods). (**e**) Extrinsic edge suppression in the presence of attention. Responses to the occluding edge are compared for attention on background versus attention on foreground. Histograms represent a suppression ratio, as symbolized by the division operator, for the neuron corresponding to the attended region.
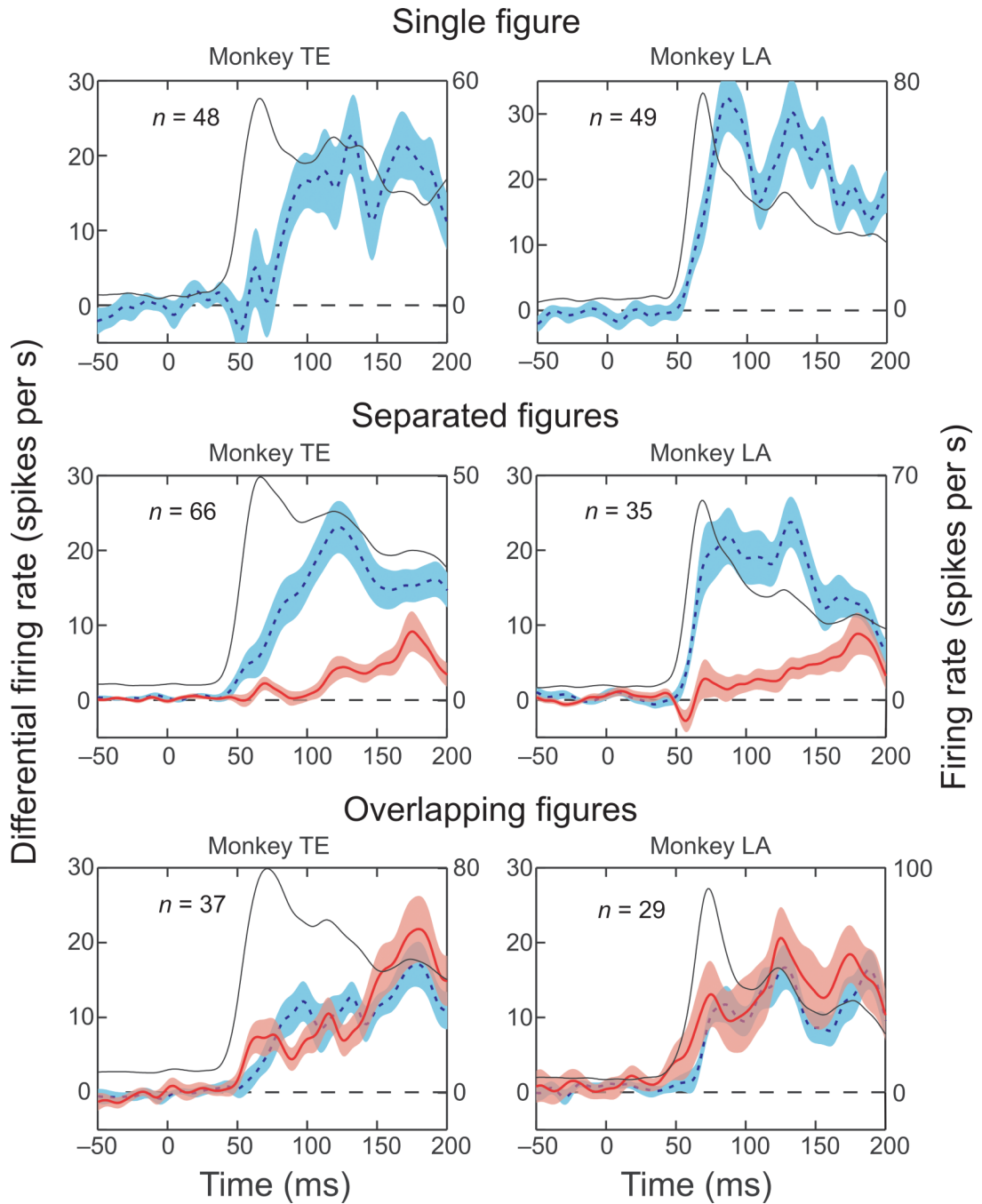
Figure 4.
The time course of border ownership signal and attention modulation. Instantaneous firing rate differences are plotted as a function of time after stimulus onset. Top, presentation of single figure in fixation task. Middle, attention task with separated figures. Bottom, attention task with overlapping figures. Dashed line + blue shading, border ownership signal; solid line + red shading, attention modulation (both left scale). Solid black line, mean response (right scale). Average over all cells with significant influence of both border ownership and attention (p<0.05, ANOVA); shading, ±1 s.e.m.. Gaussian smoothing with σ = 5ms.
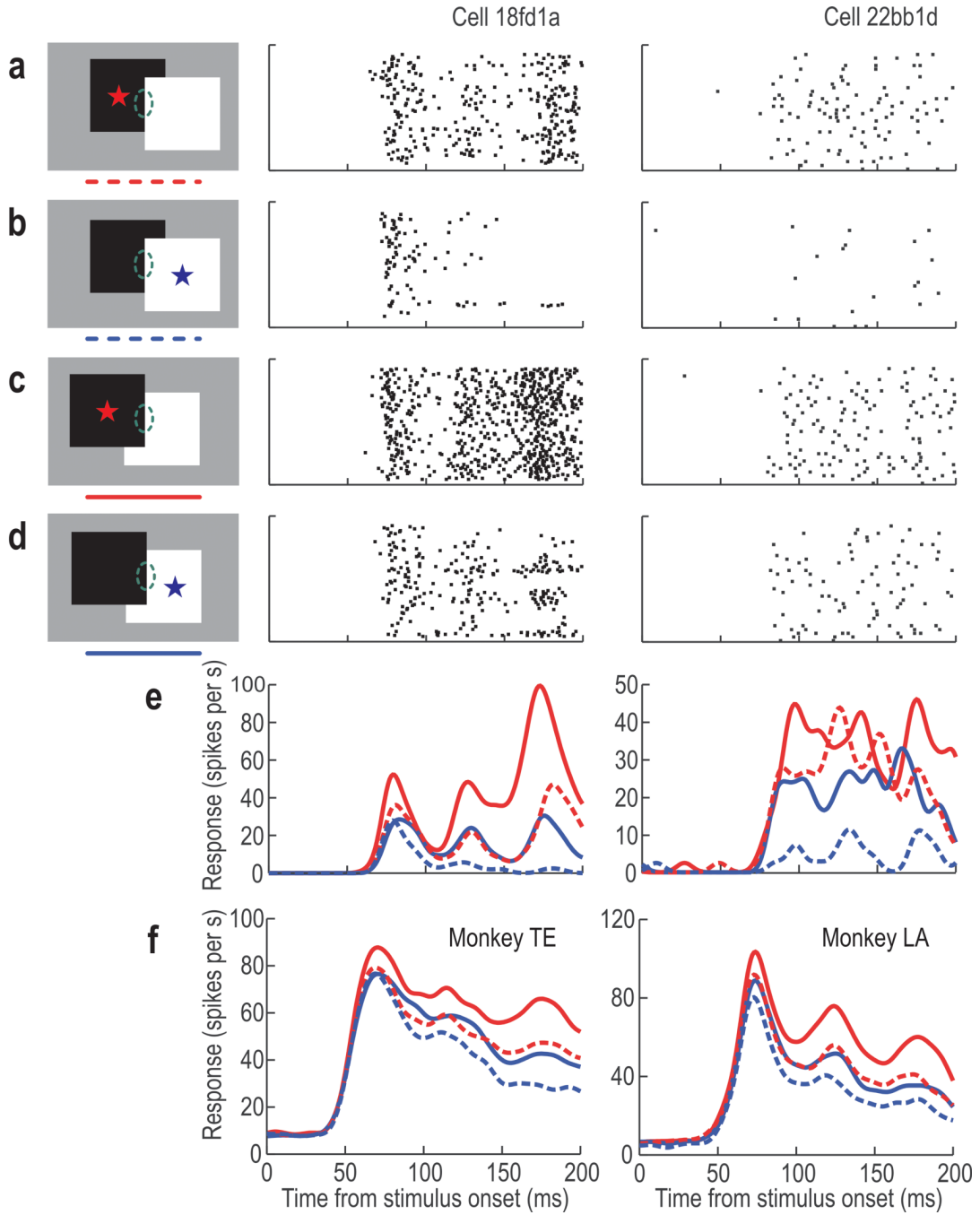
Figure 5.
The responses of two example neurons to the border between overlapping figures. Raster plots show the spike activity under four experimental conditions, as illustrated on the left. In **a–b** the right figure is in front (non-preferred side), in **c–d** the left figure is in front (preferred side). Asterisks indicate location of attention, ellipses symbolize RF. Curves at the bottom show the instantaneous firing rates as a function of time for the example neurons (**e**), and for the population of neurons influenced by border ownership and attention (main effects or interaction), for each monkey (**f**); **solid** lines: preferred border ownership, **dashed** lines: nonpreferred border ownership; **red**: attention left, **blue**: attention right. Note that attention-

left produced bigger responses than attention-right for both overlap conditions (**a** > **b** and **c** > **d**).
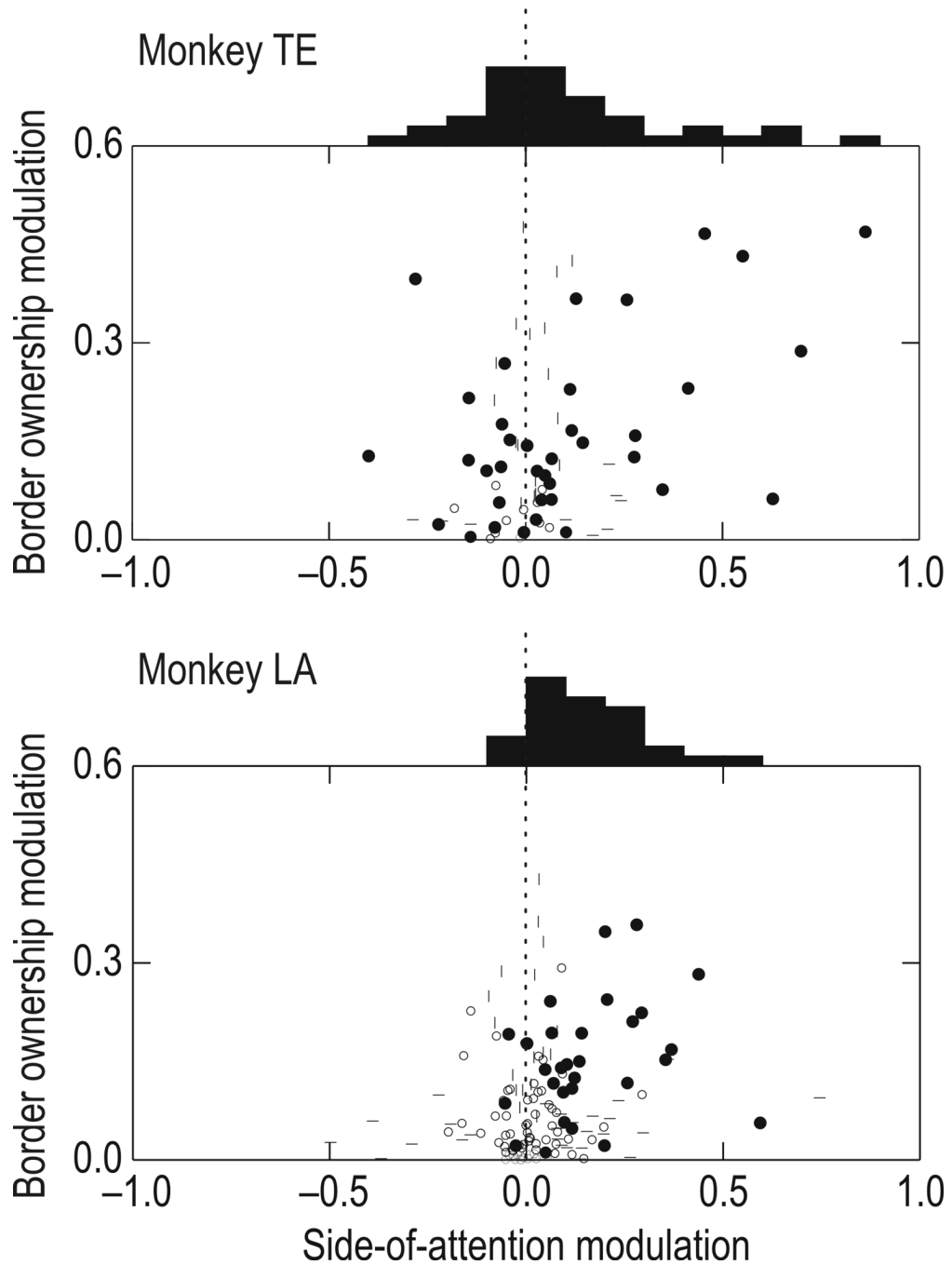
Figure 6.
Correlation between border ownership modulation and spatial asymmetry of attention effect. Responses to the border between overlapping figures. Border ownership modulation is plotted versus side-of-attention modulation. Positive side-of-attention modulation means that attention produced enhancement on the preferred side of border ownership, negative modulation means that it produced enhancement on the opposite side. Filled circles represent neurons that showed significant influences of both factors (in form of main effects or interaction); vertical dashes, border ownership selectivity without influence of attention; horizontal dashes, side-of-attention effect without influence of border ownership; open circles, no significant influence. For most cells, attention enhancement and border ownership preference are on the same side.

The histograms show the distribution of the side-of-attention modulation index in neurons influenced by both factors.
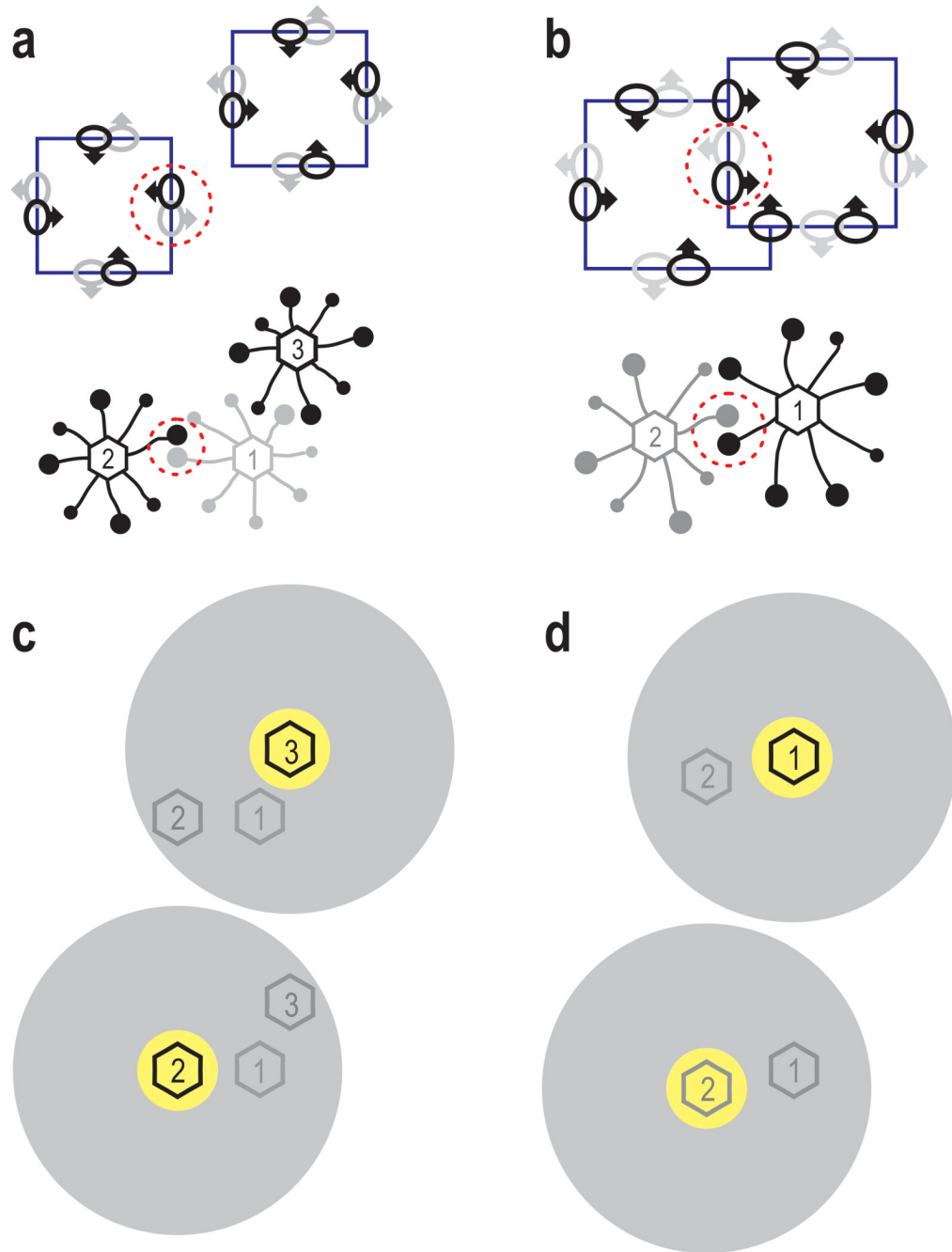
Figure 7.
A model of figure-ground organization and selective attention. (**a–b**) Top, stimuli and receptive fields, below, neural cortical circuits of border ownership selective cells (B-cells). Black disks, B-cells; large discs indicate cells stimulated by edges. Numbered hexagons, grouping cells (G-cells). G-cells integrate signals of B-cells with approximately co-circular receptive fields and, via feedback connections, set the gain of the same B-cells. Each B-cell connects to G-cells on one side only; dashed circles mark pairs of B-cells with opposite G-cell connections. Convex shapes such as squares activate G-cells in the center of their cortical representation, increasing the gain of transmission in the corresponding B-cells (border ownership modulation). (**c–d**) Volitional attention activates G-cells at the focus of attention (yellow region) or inhibits G-

cells surrounding it (gray region). Top and bottom graphs show the conditions when attention is focused on right and left figures, respectively, as required by the task in the present experiments. Because G-cells modulate the gain of B-cells, top-down attention enhances and suppresses edge signals accordingly.