Neural representation of transparent overlay

Fangtu T Qiu & Rüdiger von der Heydt

Perceptual transparency is a surprising phenomenon in which a number of regions of different shades organize into overlaying transparent objects. We recorded single neuron responses from *Macaca mulatta* area V2 to a display of two bright and two dark squares that appeared as two overlaying bars. We found that neurons assign border ownership according to the transparent interpretation, representing the shapes of the bars rather than the squares.

Perceptual transparency has been studied extensively[1–3], but the underlying neural mechanisms have rarely been investigated[4], presumably because of the uncertainty about how transparency is coded in the neural activity. A tool for studying transparency is provided by recent work on figure-ground coding in which neurons in the visual cortex were found to signal the assignment of borders to regions ('border ownership')[5,6]. Perception of transparency is linked to border ownership assignment. For example, the border marked by an ellipse in **Figure 1a** is part of the contour of the small square. However, in **Figure 1b** it is perceived as the edge of a transparent vertical bar. The light and dark regions abutting within the ellipse are exactly the same, but their interpretation by the visual system has changed. When the corners of the light and dark squares are rounded off (**Fig. 1c**), transparency is no longer perceived, and the marked edge again appears as the contour of the small square. The assignment of the edge to one or the other side determines which shapes are perceived. Thus, resolving transparency and assigning border ownership are fundamental operations for the interpretation of images[7,8].

In the secondary cortical area V2, a large fraction of orientation-selective neurons respond to the same local contrast border with different strength depending on whether the border is the contour of a figure on one side of the receptive field or the other[5]. The differential firing rate of opponent pairs of such neurons is thought to signal border ownership[5,6]. These findings relate to the task of figure-ground segregation, a process that assigns order to surfaces in depth. The phenomenon of transparency is different in that it does not imply depth ordering. Our transparency configuration (**Fig. 1b**) can be perceived as a shadow cast across a white object (or as a beam of light across a dark object) without implication of depth. Thus, the system does not rely on depth cues to resolve transparency. How do neurons respond to transparent overlay? We measured the responses of a V2 neuron to three different object configurations (**Fig. 1a–c**). The marked edges were presented in the receptive field (**Fig. 1**, ellipses) at the neuron's preferred orientation. Each figure was also flipped about this edge and the contrast was reversed (**Fig. 1a–c**, lower images), so that the edge in the receptive field was identical in all figures. The transparent configuration differed from the single-square display only in the addition of three squares, all of which were well outside the classical receptive field of the neuron (the center square was the same gray as the background). The neuron's response to the isolated square was strongest when the square was to the left of the receptive field (**Fig. 1a**, top). However, our transparency configuration elicited the strongest response for the flipped configuration (**Fig. 1b**, lower) in which the apparent bar was to the left, and the square to the right of the receptive field, in agreement with the perception of border ownership. When the corners of the squares were rounded off and the marked edge
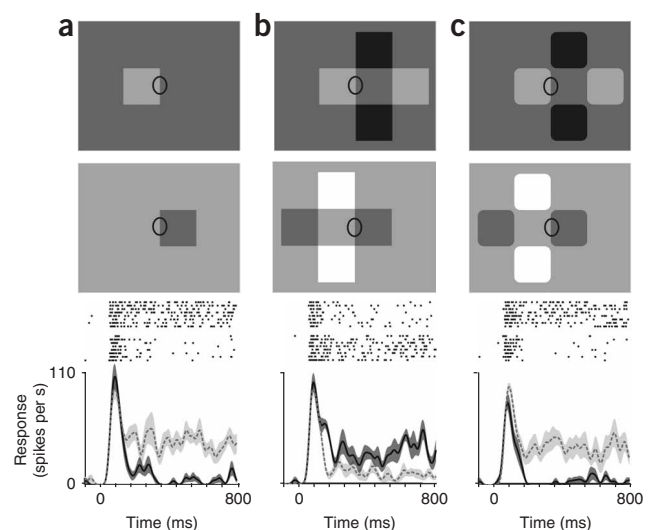
Figure 1 Neural border ownership assignment in situations of transparent overlay. (**a–c**) A V2 neuron was tested with three different configurations, one of which appears as a transparent overlay of two bars (**b**). Each configuration was applied in two mirror-image versions with local edge contrast preserved, as illustrated in the two rows of displays (ellipse marks the location of the receptive field). Below, raster plots represent corresponding responses of the neuron. The curves show the smoothed mean firing rate; shaded region indicates ± s.e.m. Dashed line and light shading correspond to the top displays; solid line and dark shading correspond to the bottom displays. Displays with the opposite contrast polarity (not illustrated) were also tested; raster plots and curves include the responses for both polarities. Typically, the squares measured 3° on a side, with luminances of 93, 62, 32 and 2 cd m^{-2} for white, light gray, dark gray and black, respectively. Single neuron responses were recorded from the visual cortex of macaques during behaviorally induced fixation. See **Supplementary Note**.

Krieger Mind/Brain Institute and Department of Neuroscience, Johns Hopkins University, 3400 N Charles Street, Baltimore, Maryland 21218, USA. Correspondence should be addressed to R.v.d.H (von.der.heydt@jhu.edu).
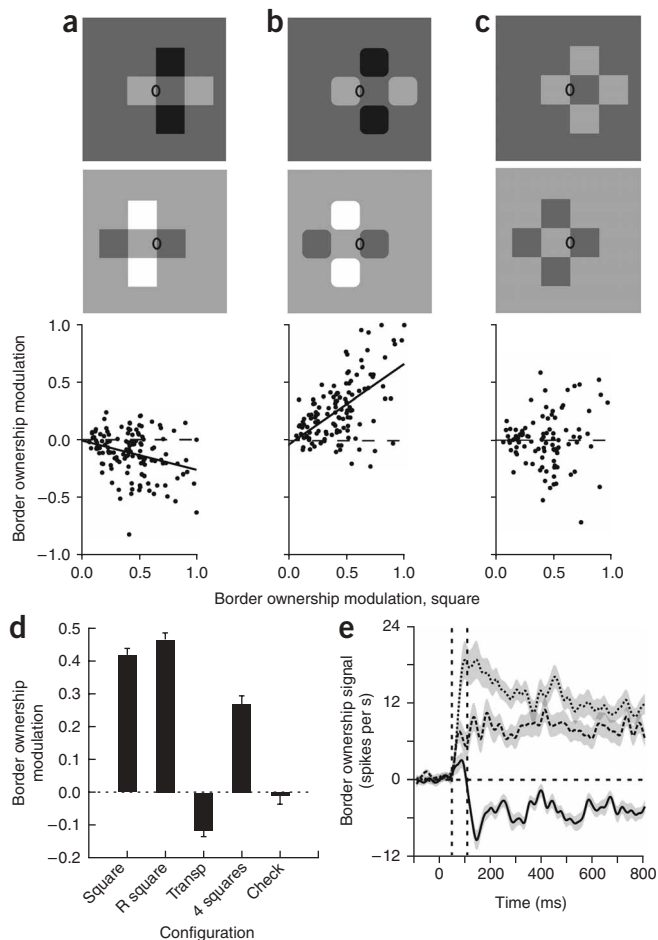
Figure 2 The influence of transparency cues on neural border ownership signals in V2. (**a**–**c**) Scatter plots of the border ownership modulation index for the illustrated configurations as a function of the index for an isolated square (**a**,**b**, $n = 127$; **c**, $n = 100$). (**d**) Mean modulation index with s.e.m. for a square, a square with rounded corners (R square) and configurations **a** (Transp), **b** (4 squares) and **c** (Check). The transparent cross produced negative modulation (that is, modulation opposite to that of a single square); the four rounded squares produced positive, but reduced, modulation; and the checkerboard produced no modulation, on average. (**e**) Time course of the mean border ownership signal for isolated square (dashed line), four rounded squares (dashed, bold line) and transparent cross (solid line) in the cells that showed reversal of modulation in the transparent condition ($n = 90$). Shaded areas indicate ± s.e.m. Note the delayed reversal of the border ownership signal in the transparent condition.

border ownership modulation, whereas the four rounded squares produced a somewhat lower modulation. The checkerboard pattern produced zero net modulation, which is consistent with the observation that, in this case, the test edge can be associated with either the left square or the central square. In the transparent condition, border ownership modulation was reversed. This reversal is notable considering the subtle difference between the configurations (**Fig. 2a,b**): the presence or absence of 'X junctions'[9]. The negative modulation for the transparent cross means that the square regions forming the arms of the cross are not represented as squares—they do not own the fourth edge—but as the ends of two crossed bars.

We determined a time course of the border ownership signal (firing-rate difference) for the single square, the control figure of four rounded squares and the transparent cross (**Fig. 2e**). The curves are averaged over all neurons that showed the signal reversal in the transparent condition. All signals showed an onset at about 50 ms. The signal for the transparent condition first deviated in the positive direction before turning negative, which occurred at 110 ms (the positive deviation was significant; $P = 0.028$, $n = 90$, two-sided t-test). Thus, the influence of the X junctions seems to build up later than the influence of the cues that determine border ownership for squares.

The neural signals paralleled perception in all of our tests. Taken together, these results are compelling evidence for mechanisms designed to resolve transparent overlay. The strong influence of transparency on the border ownership signals suggests that the underlying mechanisms in V2 serve not only figure-ground segregation, but the more general function of grouping contour features for the purpose of object identification.

Note: Supplementary information is available on the Nature Neuroscience website.

again appeared as the contour of the small square, the neuron responded most strongly when the small square was to the left of the receptive field (**Fig. 1c**, top).

To characterize the behavior of each cell, we calculated a border ownership modulation index (the normalized difference in firing rate produced by flipping the configuration, assigning the sign of the difference for each neuron according to the preferred side for the single square condition; see **Supplementary Note** online). We compared our transparency configuration (**Fig. 2a**), a configuration of squares with rounded corners (**Fig. 2b**), and a 'checkerboard' configuration (**Fig. 2c**). This last configuration included four squares of the same contrast polarity. It does not suggest transparency: it is perceived as five squares rather than two crossed bars. We plotted the modulation index for these three configurations against the modulation index for the single square. Only cells with significant modulation in the single square condition were included ($P < 0.05$, 127 of 244, **Fig. 2a,b**; 100 of 173, **Fig. 2c**). The response modulation for the transparent bars tended to be negative (90 of 127, $P = 2.5 \times 10^{-6}$, proportion test; negative modulation significant at $P < 0.05$ in 30 cells, ANOVA), and was negatively correlated with that for the single square, whereas modulation for the rounded square configuration tended to be positive and positively correlated. The index for the checkerboard pattern showed no correlation with the index for the single square.

We compared the mean values of the modulation index for five configurations: a square, a square with rounded corners, the transparent cross, the four rounded squares and the checkerboard (**Fig. 2d**). The single squares of both types produced virtually the same degree of

1. Metelli, F. *Sci. Am.* **230**, 91–98 (1974).
2. Beck, J., Prazdny, K. & Ivry, R. *Percept. Psychophys.* **35**, 407–422 (1984).
3. Adelson, E.H. *The New Cognitive Neurosciences* (ed. Gazzaniga, M.) 339–351 (MIT Press, Cambridge, Massachusetts, USA, 2000).
4. Stoner, G.R. & Albright, T.D. *Nature* **358**, 412–414 (1992).
5. Zhou, H., Friedman, H.S. & von der Heydt, R. *J. Neurosci.* **20**, 6594–6611 (2000).
6. Qiu, F.T. & von der Heydt, R. *Neuron* **47**, 155–166 (2005).
7. Nakayama, K., Shimojo, S. & Silverman, G.H. *Perception* **18**, 55–68 (1989).
8. Nakayama, K., Shimojo, S. & Ramachandran, V.S. *Perception* **19**, 497–513 (1990).
9. Adelson, E.H. & Anandan, P. *Proceedings of the American Association for Artificial Intelligence Workshop on Qualitative Vision, Boston, MA, 1990*, 77–81.

Neural Representation of Transparent Overlay

Fangtu T. Qiu & Rüdiger von der Heydt[1]

Methods

Single neuron activity was recorded extracellularly from area V2 of the visual cortex in behaving macaques (Macaca mulatta). The data are from three animals (6 hemispheres). Quartz-insulated Pt-W or epoxy-insulated tungsten microelectrodes were inserted through the dura mater within small (3-5 mm) trephinations. Area V2 was identified by its retinotopic organization and, for about half of the data, by histological reconstruction of the recording sites as described previously 1. Receptive field eccentricities ranged between 0.03 and 4.9 deg (median 1.5). Fixation was controlled by rewarding the animal for looking at a fixation spot, which was controlled by infra-red video based monocular eye movement recording (Iscan ETL-200) with a resolution of 5120 (H) and 2560 (V). One eye was imaged through a hot mirror in such a way that the camera was on the axis of the eye when the monkey fixates. We found that this design reduced signal distortion and errors due to pupil contractions, compared to the usual off-axis camera method. 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. Fixation for a minimum of 1.1 s was required. Further details of our general methods are described in Ref 2. 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.

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 either on a Barco CCID 121 FS color monitor with 1280x1024 resolution, a 72 Hz refresh rate, and a maximum luminance of 62 cd/m^2, or on a 21-inch EIZO FlexScan T965 color monitor with 1600x1200 resolution, a 100 Hz refresh rate, and a maximum luminance 93 cd/m^2. Stationary bars were used to determine the color preference, and bars and drifting gratings to map the 'minimum response field' of each cell.[1] Orientation tuning curves were recorded using moving bars. In part of the experiments, accounting for 2/3 of the data, the stimuli were presented stereoscopically[2] and disparity selectivity was also assessed. If a cell was found to be disparity selective, the optimum disparity was applied to the whole figure, relative to the fixation point, in the following tests. In most cases, however, the stimuli were viewed binocularly with zero disparity. The results regarding transparency coding were indistinguishable. The stimuli were monochromatic; generally neutral grays were used, except for neurons with pronounced color selectivity in which the preferred hue was applied.
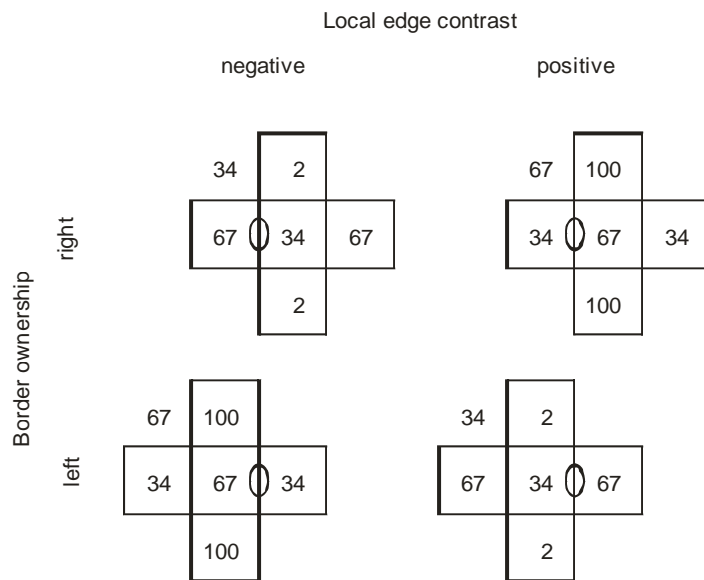

Figure S1 Distribution of the luminance values in the transparent cross stimuli. The luminances of the 5 square regions and the background regions are indicated in relative units. Ellipse indicate location of receptive field. Two displays with the same local contrast polarity, but mirror-image configurations, were compared for measuring border ownership modulation. The same was done for each local contrast polarity.

[1] Krieger Mind/Brain Institute & Dept. of Neuroscience, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218
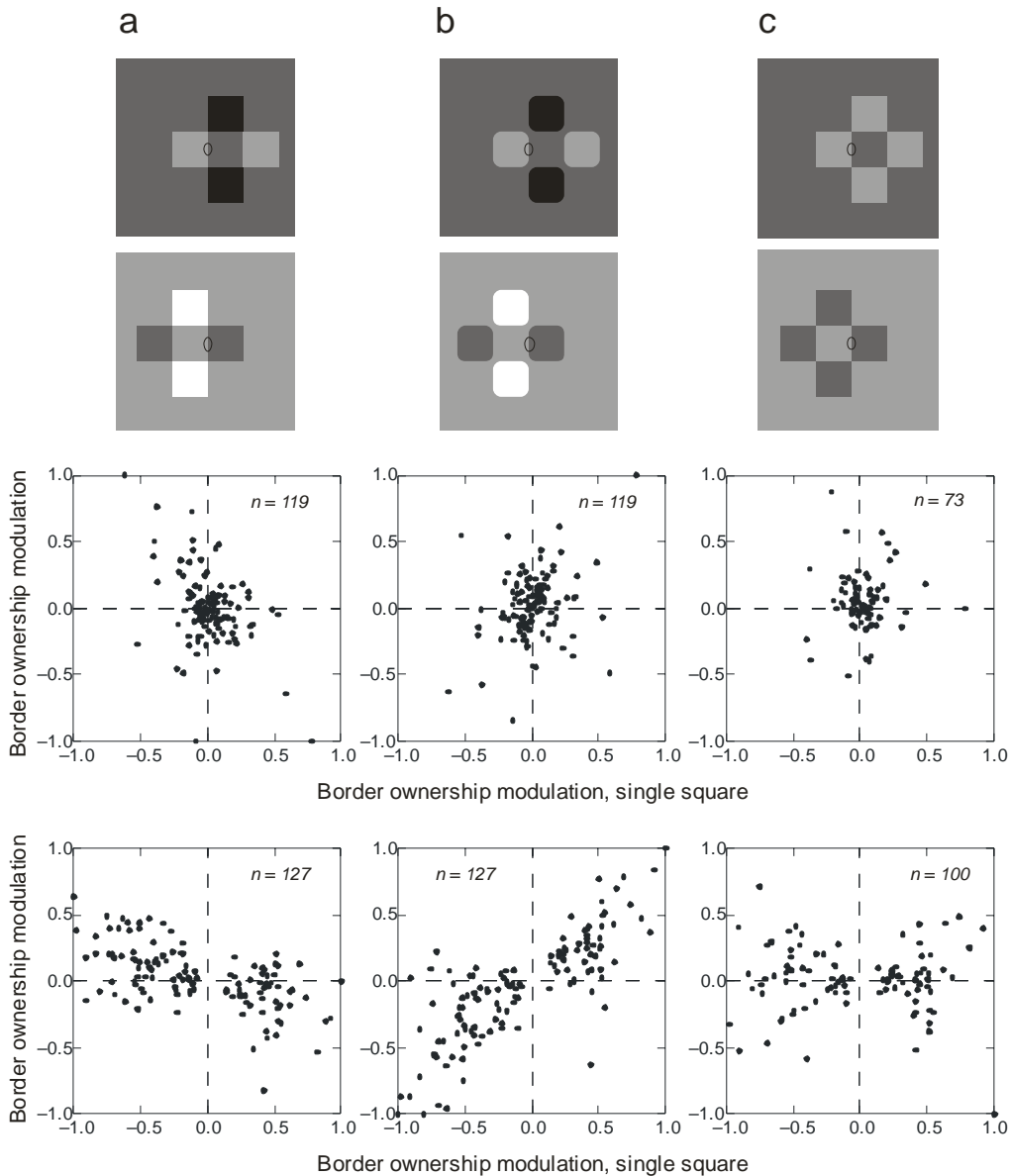
Figure S2 Summary of results from all cells tested. For these plots, the sign of border ownership modulati on was not assigned according to the side preference for the single square and modulation are thus rep resented in four quadrants. Plots in center row represent cells whose border ownership modulation did not r each the criterion of 5% significance (cells excluded in Fig. 2). Plots in bottom row show the data of Fig. 2 plotted in the same way for comparison.

After mapping the receptive field and assessing the basic stimulus preferences, the stimuli of **Figs. 1-2** were presented with the test edge centered on the minimum response field at the preferred orientation, as shown by the ellipse. The size of the individual squares was set to about twice the size of the classical receptive field. Typically the squares measured 3 degrees visual angle on a side (occasionally 2 deg or 4 deg, depending on eccentricity and receptive field size). Stimuli were usually achromatic. However, for eight cells with pronounced color selectivity chromatic stimuli were used that varied in luminance, but not in hue. **Fig. S1** shows the luminances of the different regions of the transparent cross displays relative to the maximum luminance. The luminance in the center of the cross was the same as the background luminance. Note that the X-junctions were of the 'non-reversing' type, which means that the contrast polarities of the edges were preserved across the junction.[3] The stimulus of **Fig. 1c** consisted of four squares of the same luminances and size as in the transparent cross, except that their corners were rounded (radius = 0.5 deg for 3 deg square size). The stimulus of **Fig. 2c** consisted of four squares of equal luminance (67 with background 34, or 34 with

background 67). A single square (**Fig. 1a**) and a similar square with rounded corners were also tested. Each of the five configurations was applied in four variants, for the two border ownership conditions and the two edge contrast polarities, as indicated in **Fig. S1**. Note that the squares abutting in the receptive field were identical in the two border ownership conditions. One stimulus condition was presented per fixation period, and each of the 20 conditions was presented four times, in randomized sequence. Data from fixation periods in which the eye movement monitor indicated deviations of gaze greater than 0.67 deg were discarded and the pertaining stimulus condition was repeated. Each trial started with a blank screen with the background luminance and a fixation spot in the center. After 300 ms of fixation, the stimulus pattern was presented.
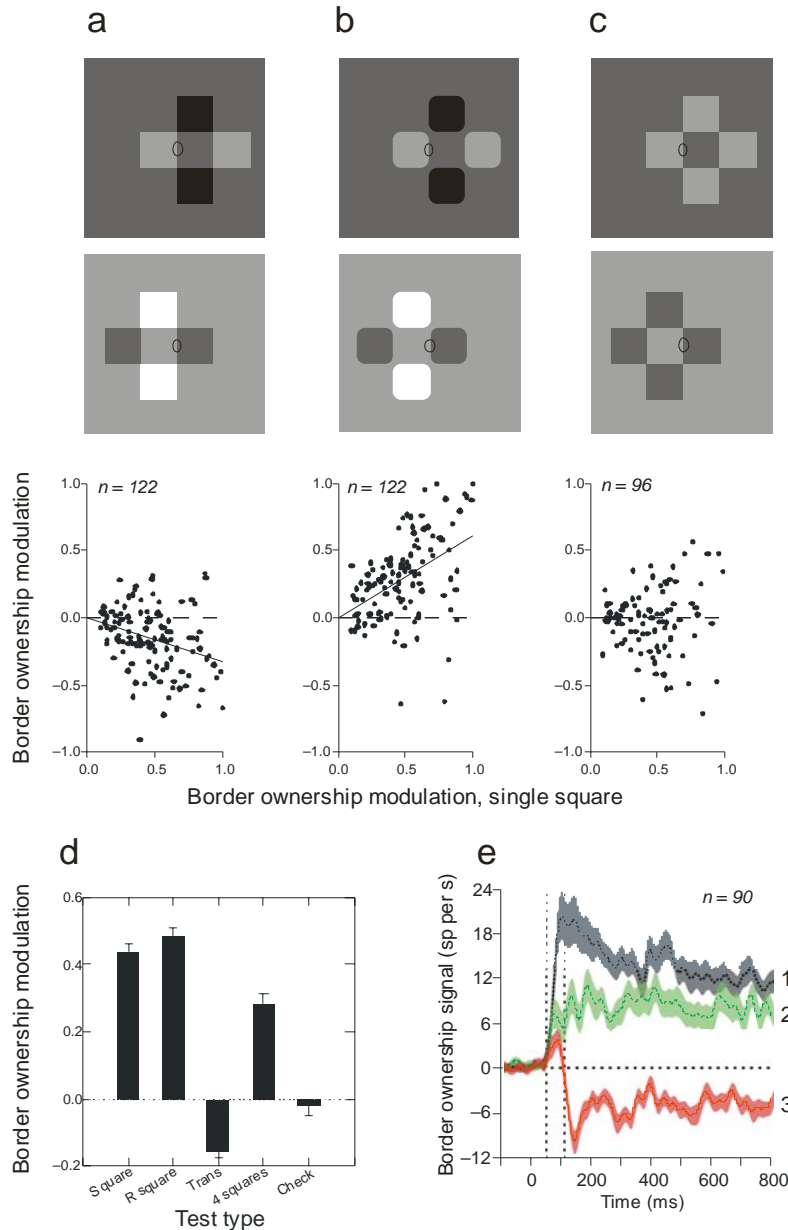

Figure S3 Results as in Fig. 2, but with a different time window for spike counting. Instead of including the entire stimulus presentation period, firing rates were calculated for the window 120-800 ms after s timulus onset. This leads to somewhat larger estimates of border ownership modulation compared to Fig. 2, particularly for the transparent condition. Other slight differences are due to different selectio n of cells according to significance of border ownership effect.

The border ownership signal was defined as the difference between the mean response for the preferred and the mean response for the nonpreferred side (averaging over contrast polarity). The side preference was determined for the single square condition and the same assignment was used for all other conditions. Generally, border ownership selectivity was assessed, for each neuron and condition, by 2-factor analysis of variance (ANOVA), the factors being side-of-figure and edge contrast polarity. The data from the control experiment with two sizes of squares (**Fig. S4**) were analyzed with 3-way ANOVA (side-of-figure, edge contrast polarity, and size). The significance criterion was 0.05. The analysis was based on the spike counts during 800ms after stimulus onset, which were transformed according to $\sqrt{count + 0.4}$ to equalize the variances. A border ownership modulation index was calculated using the conventional formula for the modulation index on the back-transformed means:

$$Index = \frac{[E(\sqrt{count_1 + 0.4})]^2 - [E(\sqrt{count_2 + 0.4})]^2}{[E(\sqrt{count_1 + 0.4})]^2 + [E(\sqrt{count_2 + 0.4})]^2 - 0.8}$$

where E(x) indicates the mean of x. The spontaneous activity was not subtracted from the responses.
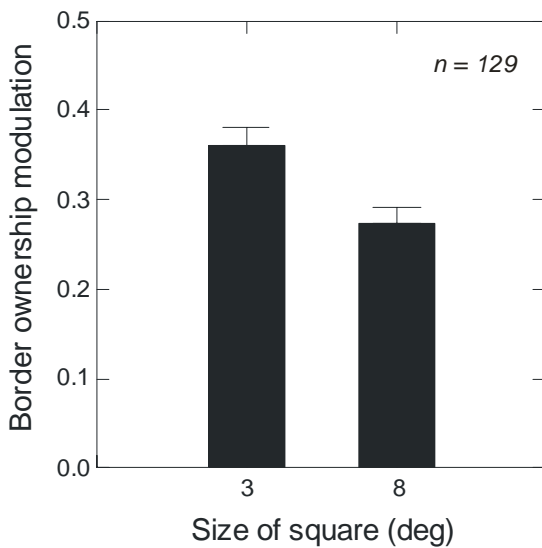

Figure S4 Comparison of the border ownership modulation in response to squares of different side length, 3 deg and 8 deg. The larger square produced 25% smaller modulation, suggesting that solid bars of the size and shape as those of the transparent cross (typically 9 x 3 deg) would also produce smaller modulation than the square that was used for comparison in Fig. 2.

Notes

Influence of fixational eye movements

Due to the design of the experiments, variations of fixation could produce results similar to those we have obtained only under special circumstances. Consider first the possibility that gaze position varied between stimulus conditions **a-b** of **Fig. 1**, but was constant within conditions. Because the border ownership signal is the difference between the responses to two stimuli that were identical over a region of 6 x 3 deg centered over the response field (see insets in **Figs. 1-2**), and the response fields typically measured only 1-2 deg in diameter, one would have to postulate that the stimulus edge was de-centered by more than 0.5-1 deg (despite our efforts to center it) so that the remote edges of the figures entered the response field to produce a spurious border ownership difference. For the square condition this would have to occur in a majority of cells (**Fig. 2**). A systematic shift of gaze position between the 'square' and the 'transparent' conditions would then have to be postulated to explain a reduction, and even reversal, of the border ownership difference.

We first compared the mean positions of gaze between the transparency condition and the four other conditions, averaging over all fixation periods that contributed to the recordings of the cells represented in **Fig. 2**. The result showed all eight position differences (4 pairs of conditions, horizontal and vertical) were smaller than 0.008 deg in absolute value (N=2703 for 'square' and '4 squares'; N=1952 for 'rounded square' and 'checkerboard'). Thus, the mean fixation position did not change between conditions. It is possible, however, that shifts of fixation did occur, but their direction depended on the stimulus orientation, and would thus cancel in the average (because the orientation varied from cell to cell). We have therefore transformed the eye movements into a coordinate system relative to the preferred orientation

and preferred direction of border ownership of each neuron, and analyzed both components of eye movements in these neuron specific coordinates. Again, absolute values of the mean differences between conditions were all smaller than 0.0084 deg (average 0.0026), and none was significant.

Another way to argue would be to assume that border ownership differences are produced by smaller eye movements that depend also on the side of the figure. For example, eye movements might be induced by the flipping of the figure to one or the other side of the receptive field, which occurred randomly between trials. This, in combination with eccentric positioning, would produce spurious border ownership response differences by displacing the RF perpendicular to the edge. Such eye movements would then have to be different between the 'square' and 'transparent' conditions to produce the results of **Fig. 2**. However, we have previously shown[1] that flipping a square between trials does not induce eye movements, and a similar analysis of the present data showed the same result.

Thus, contributions of eye movements to the results shown in **Fig. 2** can be ruled out.

Summary of results including all cells

For a complete representation of our data we prepared scatter plots similar to those of **Fig. 2** also for the cells that did not have a significant border ownership effect (**Fig. S2**, center row). Because we cannot reasonably assign preferred sides in these neurons we have plotted the border ownership modulation in spatial coordinates (positive meaning figure left, disregarding the vertical dimension). In the bottom row, the data of **Fig. 2** have been replotted in the same way. We found small, but significant correlations with the single square modulation for the transparent cross (Spearman rank correlation coefficient -0.24, N=119, p=0.01), and for the rounded squares (+0.22, N=119, p=0.014), but not for the checkerboard (-0.01, N=73, p=0.80). Thus, like the cells with significant border ownership effect, the remaining cells tended to produce the opposite modulation in the transparent condition as in the single square condition, the same direction of modulation for the control stimulus of rounded squares, and no modulation for the checkerboard pattern.

The amplitude of border ownership modulation

The analysis presented in **Fig. 2** was based on average firing rates during 800 ms of fixation after stimulus onset. This analysis tends to underestimate the amount of border ownership modulation, because unspecific activity at the beginning of response may be included in the calculation. **Fig. S3** represents the results obtained with an integration window from 120-800ms. As can be seen from the time course of the border ownership signal (**Fig. 2c**), this is approximately the optimal window. The result is a general increase of border ownership modulation, particularly noticeable in the transparent condition, compared to **Fig. 2**. The response modulation for the transparent bars was negative in 90/122 cells (significant in 39). Optimizing the response window relative to stimulus onset implicitly assumes that subsequent stages know the exact time of stimulus onset. Under natural conditions of vision that might not be the case.

Note that the comparison between conditions is biased in favor of the single square conditions (regular and rounded) because these were used to select border ownership selective cells and to assign preferred side. This bias may have resulted in an overestimation of the modulation for the squares by 10-20%, compared to the other conditions.

It could also be argued that border ownership modulation in the transparent condition should not be compared with the modulation obtained with the square, as we do in **Fig. 2**, but with the modulation that would be produced by a bar of the same size and shape as the vertical bar perceived in the transparent cross, which has a height three times that of the square. Because border ownership modulation tends to decrease with figure size[1,2], modulation to a bar of that size is likely to be smaller than the modulation obtained for the square. We have not recorded responses to bars of the height of the cross; however, we have additional data from a test in which squares measuring 3 deg and 8 deg on a side were presented with either contrast polarity and on either side of the receptive field. This test was generally run first after the receptive field of a neuron and its basic response properties were determined. The comparison between border ownership modulation for the two sizes of squares can provide a clue as to how big border ownership modulation would be for a bar of 9 deg height (the most frequently used height of the transparent cross). **Fig. S4** shows the average border ownership modulations for the cells that showed a significant effect (n=129). It can be seen that the modulation for the larger square was about 25% smaller than that for the smaller square. Since the 8 deg square is only slightly shorter than the 9 x 3 bar, but much wider, it should produce less modulation than the bar. Thus, the modulation for the bar would likely be in between those for the 3 deg and the 8 deg squares, that is, between 100% and 75% of the modulation obtained with the square in the main experiment.

We chose to compare the transparent cross with the square because in these two configurations stimulation is identical in a 6 x 3 deg region around the receptive field, as well as in the background surrounding the figures (see **Fig. S1**); the transparent stimulus differed from the single square only in the addition of three squares outside the RF. From

the diagram in **Fig. S1** it can be seen that, if one uses a bar the size of the apparent vertical bar for comparison, the condition of identical local stimulation requires that the bar has the same luminance as the center square of the transparent cross, which means that the background luminance must be different. Changes in responses could then be attributed to the change in background luminance.

References

1. Zhou,H., Friedman,H.S., & von der Heydt,R. Coding of border ownership in monkey visual cortex. *J. Neurosci.* **20**, 6594-6611 (2000).
2. Qiu,F.T. & von der Heydt,R. Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. *Neuron* **47**, 155-166 (2005).
3. Adelson,E.H. & Anandan,P. Ordinal characteristics of transparency. *AAAI-90 Workshop on Qualitative Vision* 77-81 (1990).