Systems/Circuits

Consistency of Border-Ownership Cells across Artificial Stimuli, Natural Stimuli, and Stimuli with Ambiguous Contours

Janis K. Hesse[1,2] and Doris Y. Tsao[1,2]

[1]Howard Hughes Medical Institute, Division of Biology and Biological Engineering and [2]Computation and Neural Systems, California Institute of Technology, Pasadena, California 91125

Segmentation and recognition of objects in a visual scene are two problems that are hard to solve separately from each other. When segmenting an ambiguous scene, it is helpful to already know the present objects and their shapes. However, for recognizing an object in clutter, one would like to consider its isolated segment alone to avoid confounds from features of other objects. Border-ownership cells (Zhou et al., 2000) appear to play an important role in segmentation, as they signal the side-of-figure of artificial stimuli. The present work explores the role of border-ownership cells in dorsal macaque visual areas V2 and V3 in the segmentation of natural object stimuli and locally ambiguous stimuli. We report two major results. First, compared with previous estimates, we found a smaller percentage of cells that were consistent across artificial stimuli used previously. Second, we found that the average response of those neurons that did respond consistently to the side-of-figure of artificial stimuli also consistently signaled, as a population, the side-of-figure for borders of single faces, occluding faces and, with higher latencies, even stimuli with illusory contours, such as Mooney faces and natural faces completely missing local edge information. In contrast, the local edge or the outlines of the face alone could not always evoke a significant border-ownership signal. Our results underscore that border ownership is coded by a population of cells, and indicate that these cells integrate a variety of cues, including low-level features and global object context, to compute the segmentation of the scene.

Key words: border-ownership; feedback; figure-ground; illusory contours; natural; segmentation

---

Significance Statement

To distinguish different objects in a natural scene, the brain must segment the image into regions corresponding to objects. The so-called "border-ownership" cells appear to be dedicated to this task, as they signal for a given edge on which side the object is that owns it. Here, we report that individual border-ownership cells are unreliable when tested across a battery of artificial stimuli used previously but can signal border-ownership consistently as a population. We show that these border-ownership population signals are also suited for signaling border-ownership for natural objects and at longer latency, even for stimuli without local edge information. Our results suggest that border-ownership cells integrate both local, low-level and global, high-level cues to segment the scene.

Introduction

The two most important tasks of the ventral stream of visual cortex are arguably segmentation of the visual scene and object

recognition. Segmentation tells us which groups of pixels in a scene constitute the fundamental units that we can interact with and recognition gives these units a meaning by telling us what they are. Often, vision is considered as a sequence of processing steps in a feedforward hierarchy (Marr, 1982) where recognition of objects happens after a series of nonlinear operations on the input image (Riesenhuber and Poggio, 2002). Because segmentation, which is thought to happen in retinotopic cortex, is earlier in the feedforward hierarchy of visual areas, it is often assumed to be a necessary step to be completed before recognition, so that the segmented regions corresponding to the object surfaces can be fed to inferotemporal cortex (IT) and recognized individually (Rubin, 1958; Nakayama et al., 1995; Driver and Baylis, 1996).
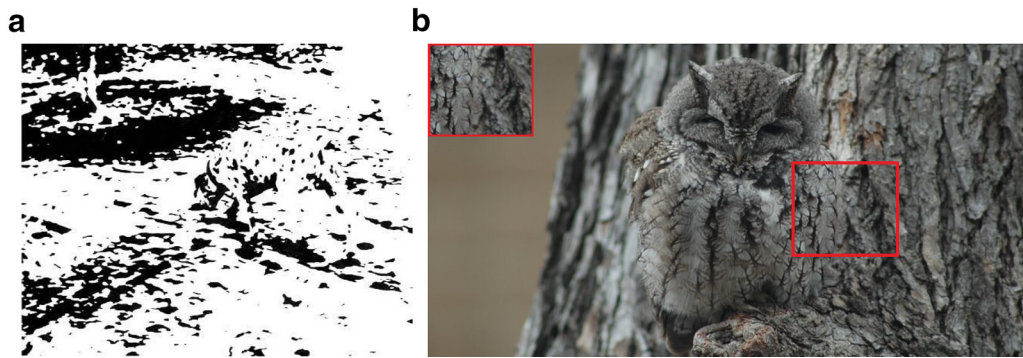
Figure 1. Segmentation can be ambiguous based on low-level cues. **a**, While the famous Dalmatian dog display is considerably more difficult to segment without knowing that there is a dog present, once one recognizes the dog, one also perceives it as a contiguous surface. **b**, The camouflaged owl seems impossible to segment based on local cues (red square, inset); but as one recognizes the owl, one can infer and perceive the boundary.

On the other hand, psychophysical studies (Peterson and Gibson, 1993, 1994; Peterson and Kim, 2001; Grill-Spector and Kanwisher, 2005) have suggested that object recognition influences or even precedes segmentation: For example, Peterson and Gibson (1993) found that recognition of object shape can overwrite depth order defined by disparity. Moreover, observers asked to report their first perceived figure-ground organization are influenced by symmetry and orientation-dependent object recognition processes and are more likely to perceive regions as figure compared with ground if they match their object memory (Peterson and Kim, 2001).

One of the most insightful neurophysiological findings for understanding how the brain segments the visual scene are the remarkable border-ownership cells discovered by von der Heydt and colleagues (Zhou et al., 2000). Border-ownership cells are thought to be crucial for segmentation as their responses signal the side-of-figure for a number of artificial stimuli. However, it is not entirely clear whether and how the side-of-figure signal observed in border-ownership cells aids recognition of objects in natural scenes. Conversely, it is not known whether object recognition in IT can influence border-ownership signals in retinotopic cortex.

Like most cells in early visual cortex, a border-ownership cell will respond to an edge presented at a given orientation in its receptive field, but it responds differentially depending on the side of the figure to which the edge belongs: A vertical edge can be the border of a foreground object that is either to the left or to the right of it. shapes Zhou et al. (2000) showed that, for a variety of artificial stimuli, border-ownership cells respond consistently more strongly if the edge belongs to a figure on its preferred side than its nonpreferred side, even if the stimuli are locally identical within the receptive field of the cell. Artificial stimuli that were previously shown to evoke consistent border-ownership responses include single luminance squares, occluding luminance rectangles, single and occluding outlines of rectangles, C-shapes (Zhou et al., 2000), disparity-defined squares (Qiu and von der Heydt, 2005), and squares evoking the percept of transparent overlay (Qiu and von der Heydt, 2007) (see stimuli used in Fig. 5a–d). Here, we asked whether border-ownership cells can also infer border-ownership for natural objects, with not necessarily straight edges and inhomogeneous, possibly confounding textures. Indeed, in natural scenes, segmentation can often be ambiguous based on only local, low-level cues (McDermott, 2004). Consider, for example, the famous Dalmatian dog display or the camouflaged owl in Figure 1. It appears impossible for an algorithm that uses only local, low-level cues to infer the correct segmentation. Yet, once we recognize the Dalmatian dog, we perceive it as an object with a contiguous surface. And we are able to infer the boundary of the owl. Bottom-up, purely feedforward algorithms would likely come to the critically different, erroneous interpretation that it is a texture. Is this perceived segmentation signal for recognizable, natural objects present in the side-of-figure signal of border-ownership cells?

To answer this question, we recorded from border-ownership cells and systematically presented a battery of both artificial stimuli and natural face stimuli, as well as face stimuli with ambiguous contours, to find out how cells that respond consistently to the side-of-figure of artificial stimuli would respond to the presented natural object stimuli.

Materials and Methods

All animal procedures used in this study complied with local and National Institutes of Health guidelines. Two male rhesus macaques were implanted with MR-compatible head posts and trained to maintain fixation on a dot for a juice reward.

Targeting. Because the interest of this study lies explicitly in recording from border-ownership cells rather than an exhaustive analysis of V2/V3, our electrode targeting was guided by fMRI. Monkeys were scanned in a 3T TIM (Siemens) magnet. Scanning procedures were the same as described previously (Tsao et al., 2006; Freiwald and Tsao, 2010; Ohayon and Tsao, 2012). For fMRI, monkeys passively viewed stimuli on a screen. MION contrast agent was injected to improve signal-to-noise ratio. To identify V2 and V3, we first mapped retinotopy by presenting horizontal and vertical checkerboard wedges and defined area boundaries based on horizontal and vertical meridians. Within V2 and V3, we targeted areas with high functional activation in response to border-rich disparity-defined checkerboard stimulus versus a full field changing disparity stimulus (Tsao et al., 2003) to increase the yield of recorded border-ownership cells (see Fig. 2a) since Qiu and von der Heydt (2005) had previously found that a majority of border-ownership cells are also selective to the side-of-figure of stereo-defined edges, which were abundant in the former stimulus. In addition, we were guided by anatomical landmarks and targets were confined to the banks and fundus of the lunate sulcus. For Monkey J, we did not get good signal in the fMRI and therefore targeted the same anatomical locations that yielded high activation in Monkey T. We found slightly less border-ownership cells than in Monkey T (see Results). Placement of recording chambers and electrode trajectories toward the targeted regions were planned with the software Planner (Ohayon and Tsao, 2012). In Monkey T, we recorded from the right hemisphere and in Monkey J from the left hemisphere.

Fixation. Monkeys were head fixed and passively viewed a screen in a dark room. A small fixation spot (0.25° in diameter) was presented in the center of the screen, and eye position was monitored using an

Table 1. Consistency of cells for border-ownership stimuli[a]

Paradigm	Significantly consistent		Significantly inconsistent		Not significant	
	All cells	Top 50	All cells	Top 50	All cells	Top 50
Standard test (single luminance square)	55% (111/201)	84% (42/50)	NA	NA	44% (90/201)	16% (8/50)
Occluding squares	19% (21/110)	30% (13/42)	10% (11/110)	2% (1/42)	71% (78/110)	67% (28/42)
Occluding outlines	11% (12/111)	17% (7/42)	16% (18/111)	5% (2/42)	73% (81/111)	79% (33/42)
C-shapes	11% (12/108)	19% (8/42)	22% (24/108)	10% (4/42)	67% (72/108)	71% (30/42)
Transparent	20% (21/107)	26% (11/42)	5% (5/107)	2% (1/42)	76% (81/107)	71% (30/42)
Four squares control	37% (40/107)	55% (23/42)	10% (11/107)	0% (0/42)	52% (56/107)	45% (19/42)
Single full faces	41% (45/111)	69% (29/42)	17% (19/111)	7% (3/42)	42% (47/111)	24% (10/42)
Occluding full faces	14% (16/111)	29% (12/42)	5% (6/111)	2% (1/42)	80% (89/111)	69% (29/42)
Single ambiguous face	20% (22/111)	29% (12/42)	10% (11/111)	0% (0/42)	70% (78/111)	71% (30/42)

[a]For different stimuli presented during the experiment, we computed the percentage of cells that were significantly tuned to the side-of-figure of the stimulus and consistent with the preferred side-of-figure for the single luminance square. Significance criterion was $p < 0.01$ as in Zhou et al. (2000). For the single luminance square (first row), the proportion of significantly tuned cells merely indicates the proportion among all 201 analyzed cells, and among the 50 cells with highest average modulation index across artificial stimuli, respectively, that were significantly tuned the side-of-figure of the single luminance square by two-way ANOVA on side-of-figure and contrast polarity. For the remaining stimuli, we only considered cells that were significantly tuned to the side-of-figure of the single luminance square, which was the case for 111 of all 201 cells and 42 of the top 50 cells, and computed the proportions of cells from these two populations that were (1) significantly tuned to the side-of-figure of the stimulus and consistent with the side-of-figure preference for the single luminance square; (2) significantly tuned to the side-of-figure but inconsistent with the single luminance square; or (3) not significantly tuned to the side-of-figure of the given stimulus. Four of the 111 cells were lost too early to present all artificial stimuli. For comparison, Zhou et al. (2000) found that, among the cells that were significantly tuned the side-of-figure of the single luminance square, 20 of 42 cells were tuned significantly and consistent to the side-of-figure of occluding squares, 1 of 42 cells was tuned significantly but inconsistent, and 21 of 42 cells were not significantly tuned. For the C-shapes, Zhou et al. (2000) found 4 of 16 cells to be tuned significantly and consistent, 12 of 16 to be not significantly tuned, and no cells to be significantly tuned and inconsistent. For the transparent stimulus, Qiu and von der Heydt (2007) found 30 of 127 cells to be tuned significantly and consistent. The significance criterion was $p < 0.05$ in Qiu and von der Heydt (2007) and $p < 0.01$ in Zhou et al. (2000).

iScan system. Monkeys were rewarded with juice for maintaining fixation every 2–3 s.

Electrophysiology. Tungsten electrodes (FHC) with 1 MΩ impedance were used for recording. Custom grids were printed, and guide tubes were cut to extend 2 mm below the dura. Electrodes were advanced using an oil hydraulic microdrive (Narishige). Neural signals were recorded using a MAP system (Plexon). Local field potentials were filtered at 0.7–300 Hz, and single units and multiunits were filtered at 0.15–8 kHz and recorded at 40 kHz.

Online data analysis. Spikes were isolated and sorted online using the box method of the SortClient (Plexon). Initially, approximate receptive field location was determined by manually sweeping a small blinking square (0.2°) across the screen. Based on this approximate location, receptive fields were mapped by computing the spike-triggered average (STA) in response to a random stimulus of size 8° that was centered on the hand-mapped location (Pack et al., 2003). The random stimulus was a series of images alternating at 100 ms that consisted of a gray background and two squares of size 0.5° appearing at random positions, with one of the squares being white and the other square randomly chosen to be either black or white. Subsequently, a 2D Gaussian was fitted to the STA of the stimulus to determine the position and size of subsequent stimuli. Receptive field maps were also computed by considering either the only black squares or the white squares alone and yielded similar receptive fields. Next, moving sine wave gratings were presented, and the preferred orientation of the cell was determined based on the sine grating orientation that evoked the highest response. For all subsequent stimuli, the central edge (indicated by a purple ellipse in the figures) was adjusted to the position and size of the receptive field and rotated to match the preferred orientation. We recorded a total of 545 cells in Monkey T and 121 cells in Monkey J. We subjectively assessed for each recorded site whether it contained promising border-ownership cells based on clear receptive fields, clear orientation preferences, and consistent responses to the side-of figure of luminance squares and for a total of 298 of 545 recorded cells proceeded to present a battery of border-ownership tests consisting of both artificial stimuli and natural stimuli of faces (some of these cells were included because they were recorded simultaneously as ones that passed the subjective assessment but were not themselves subjectively assessed). In the offline analysis, further cells were excluded based on unclear STAs or insufficient samples of responses, yielding a total of 201 valid cells (see Offline data analysis). Artificial stimuli consisted of stimuli that had been shown by von der Heydt and colleagues to evoke consistent border-ownership signals, including the standard test with single luminance squares (Zhou et al., 2000), two occluding squares (Zhou et al., 2000), occluding outlines, C-shapes, and squares that evoke the perception of either four single squares or a transparent overlay (Qiu et al., 2007). To further verify the correct mapping of the receptive field and test position invariance within the receptive field, we also performed

position tests by sliding the luminance square across 11 positions orthogonal to its preferred orientation. Natural stimuli consisted of single faces, overlapping faces, the isolated local edge of overlapping faces alone, faces with local edge deleted, Mooney faces, outlines of overlapping faces, and faces occluding apples. The whole battery of artificial and natural stimuli is shown in Figure 5. We chose to use mostly faces for the natural stimuli as they represent a natural, complex, high-level object category that is of strong behavioral and social relevance and with which monkeys have extensive experience. Also, the existence of face-selective regions in IT opens up the possibility to examine interactions between object representations in retinotopic and IT cortex (Tsao et al., 2008). Stimuli were presented for 500 ms ON time and 150 ms OFF time. To correct for delays of the screen, we used a photodiode that detected the onset and offset of the stimuli. The photodiode's output was fed into the recording system and later used to synchronize the onset of the stimulus and the neurophysiological data during offline analysis.

Offline data analysis. Spikes were re-sorted offline using OfflineSorter (Plexon). Trials in which monkeys broke fixation were discarded (using a 1° eccentricity fixation window). We discarded cells with insufficient number of trials or spikes (<500 or <1500 total spikes for standard test and natural stimulus set, respectively) and cells that either had an unclear STA or were not centered on the receptive field or failed the position test (i.e., when shifting the stimulus as in Fig. 2e, the peak response was not inside the receptive field), leaving a total of 201 cells. Peristimulus time histograms (PSTHs) were smoothed with a Gaussian kernel. For Table 1, which shows the consistency of cells for artificial and natural stimuli shown in Figure 5, we determined the proportion of side-of-figure selective cells for different stimuli using the same method as by Zhou et al. (2000): Side-of-figure selectivity of a cell to a given stimulus was computed using a two-way ANOVA on side-of-figure and contrast polarity on the average firing rate from 0 to 500 ms during trials of different conditions and using an unpaired t test in case of only one contrast polarity. Unless stated otherwise, a cell was deemed significantly selective to the side-of-figure of a stimulus if $p < 0.01$ and consistent if the modulation index for the given stimulus had the same sign as for the standard test of single luminance squares. Modulation indices for pairs of matched stimuli were computed as $M_{s1,s2} = \dfrac{R_{s1} - R_{s2}}{R_{s1} + R_{s2}}$, where $R_{stimulus}$ is the firing rate of the cell in response to the stimulus averaged over the 500 ms from onset to offset of stimulus presentation. We defined the preferred side-of-figure of each cell based on the average modulation index for the standard luminance squares shown in Figure 2. Based on this preferred side, we computed the average modulation index across all pairs of matched artificial stimuli (Fig. 5a–d, blue and red conditions). This average modulation index is positive if the border-ownership coding across artificial stimuli is consistent with the border-ownership selectivity for
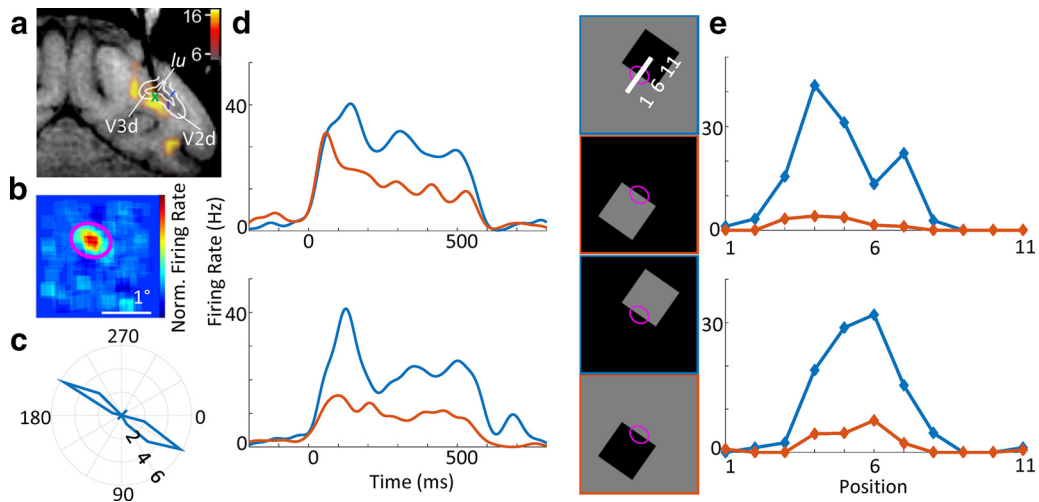
Figure 2. An example border-ownership cell. ***a***, Electrode targeting V3d; fMRI activation for disparity checkerboard versus full field disparity is overlaid. Green cross represents location of electrode tip. Dotted blue lines indicate retinotopically defined area boundaries of V2d and V3d. lu, Lunate sulcus. ***b***, Receptive field as computed by the STA. ***c***, Orientation tuning of the cell. Radius and angle of the polar plot correspond to firing rate and presented orientation of moving sine grating, respectively. This example cell had a preferred orientation of ∼30°. ***d***, PSTHs of responses to luminance squares (presented for 500 ms). Four square stimuli with different contrasts and different sides (right panel) with the edge on the receptive field (purple ellipse). This cell's response was increased when the figure was on the top right side. ***e***, Position test of the cell. To test the robustness of the border-ownership signal across positions within the receptive field, the stimulus was swept across different positions orthogonal to its preferred orientation (*x*-axis, indicated by the positions 1–11 at the top right of ***d***). Across all positions within the receptive field, the response (*y*-axis) was consistently higher when the figure was on the preferred side of the receptive field. Blue and red conditions are equivalent to the stimuli used in ***d***. Error bars indicate SEM.

the simple luminance square and negative if it is inconsistent. We selected the 50 most consistent cells with the highest average modulation indices across artificial stimuli to compute the population average responses shown in Figure 5. The results did not depend qualitatively on the exact number of selected cells. Before averaging, each PSTH was normalized by the average response from 0 to 500 ms after stimulus onset across the shown stimuli. The *p* values shown on top of each PSTH were computed using two-sided paired *t* tests on the normalized average responses from 0 to 500 ms after stimulus onset across trials of the 50 neurons for the preferred versus nonpreferred side. For the scatter plots in Figure 6, we computed the average modulation index across the two local contrast conditions of the luminance square, and the average modulation index across all artificial stimuli, and computed the correlation with modulation indices for natural stimuli. Finally, we compared latencies of the border-ownership signal for full face stimuli and ambiguous face stimuli with illusory contours. Traditionally, latencies have been computed as the time when the signal first significantly exceeds baseline fluctuations (Maunsell and Gibson, 1992; Kiani et al., 2005). However, because the response amplitudes were much lower for stimuli with illusory contours, we were worried that this definition might be biased toward longer latencies for stimuli with lower signal-to-noise ratios. Thus: (1) We only included the subset of border-ownership cells in the analysis that showed significantly stronger responses to the preferred side for both the full stimulus and the stimulus with illusory contour (*p* < 0.05, Welch test). The preferred side was determined by responses to the luminance squares. (2) We used a half-peak measurement as latency (Zhou et al., 2000): We computed the difference of smoothed (SD: 9 ms) PSTHs for preferred and nonpreferred side and defined latency as the first time that the difference reached half of the maximum difference across the 500 ms of the stimulus ON time. (3) We repeated the analysis with a change point measure (Sugihara et al., 2011), which fits a piecewise linear function consisting of two lines to the cumulative difference PSTH and defines the latency as the point where the first leg, which is fixed as 0, transitions to the second leg, which qualitatively confirmed our results.

All analysis was performed using MATLAB (The MathWorks).

Results

We targeted regions in dorsal V2 and V3 that elicited strong fMRI activation in response to disparity-defined shapes versus full-field disparity. Figure 2 shows responses of a border-ownership

cell at a representative location (Fig. 2*a*, functional activation overlaid). Figure 2*b* shows the receptive field of the example cell mapped by computing the STA. All analyses below are based on a sample of a total of 201 single units (126 Monkey T, 75 Monkey J) for which we manually verified correct positioning on the receptive field and orientation tuning. This sample is biased as we were explicitly looking for border-ownership cells and skipped cells that were not promising candidates (see Materials and Methods). For Monkey J, where the fMRI signal was weak and we were just guided by anatomical locations where Monkey T had shown high functional activation, we found slightly fewer border-ownership cells (*p* = 0.04, two-sided unpaired *t* test on average modulation indices across artificial stimuli). Receptive fields were in the lower left and lower right quadrant of the visual field for Monkey T and Monkey J, respectively, and eccentricity ranged from 1° to 5°.

Initially, the primary goal of this study was to determine how border-ownership cells respond to the side-of-figure of natural stimuli. However, preliminary recordings with a variety of artificial stimuli revealed almost no cells that were consistent in their border-ownership preference across all artificial stimuli tested, leading us to perform a systematic characterization of the consistency of border-ownership cells across a large population of cells and a large battery of artificial stimuli (for the battery of artificial stimuli used, see Fig. 5*a*–*d*). Previous studies or border-ownership cells have each focused on specific subsets of artificial stimuli. Thus, a major question remains open: whether there exists a significant fraction of "true" border-ownership cells that signal the side-of-figure reliably across all types of artificial stimuli containing object borders. Moreover, there has been considerable variability in the reported proportions of consistent cells across different studies and stimuli. Zhou et al. (2000) found more that more than half of cells in V2 to be selective to the side-of-figure of single luminance squares; among those cells, 20 of 42 cells were tuned significantly and consistent to the side-of-figure of occluding squares, 1 of 42 cells was tuned significantly but inconsistent, and 21 of 42 cells were not significantly tuned.
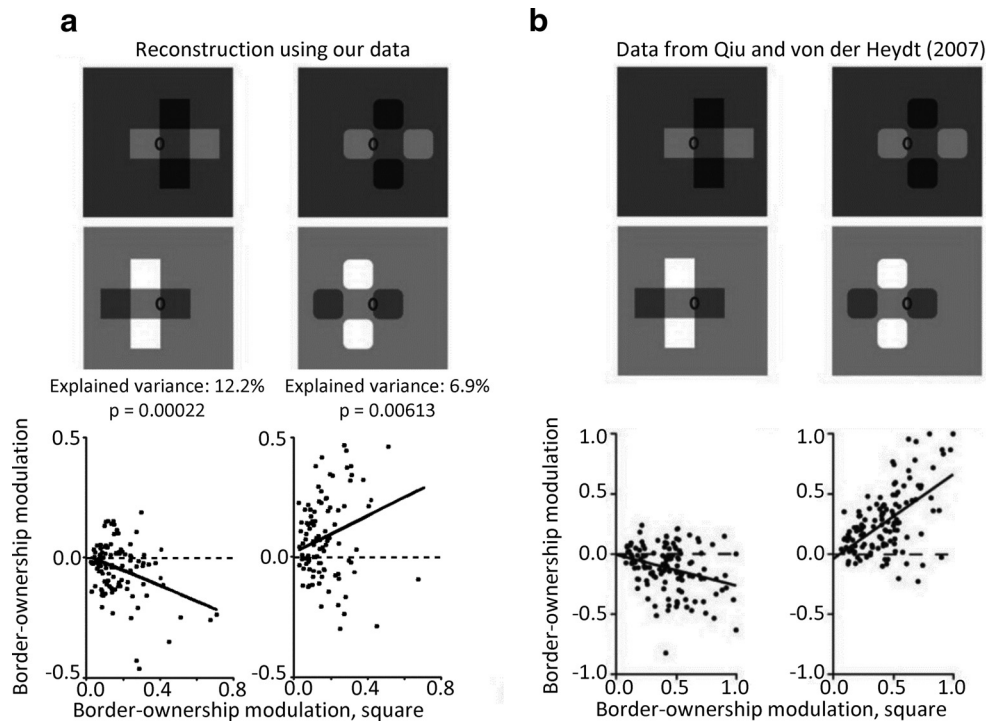
Figure 3. Comparison of population analysis with Qiu and von der Heydt (2007). We reconstructed the population analysis performed by Qiu and von der Heydt (2007) to compare border-ownership tuning of the recorded cells. **a**, Border-ownership modulation index for the single luminance square on the *x*-axis against border-ownership modulation index of the transparent overlay stimulus (left) and the four square control (right) on the *y*-axis for all cells we recorded that were significantly tuned to the side-of-figure of the single luminance square. **b**, Original data adapted from Qiu and von der Heydt (2007). For this figure, we used the same formula for modulation indices as used by Qiu and von der Heydt (2007), which is slightly more complicated than the one in the rest of this article. Thick black lines indicate Standard Model I regression lines (i.e., the ordinary least-squares fit of the abscissa to the ordinate). *p* values were computed by comparing a transform of the Pearson correlation coefficient with a Student's *t* distribution. This represents a relatively conservative measure of correlation. Because the sign of border ownership for each cell is arbitrary, we also computed correlations after duplicating all points by reflecting them over the origin. This measure yielded the same signs of correlations and even higher explained variances ($r^2 = 22.2\%$ and $p = 3.15e{-}13$ for transparent overlay stimulus, $r^2 = 24.7\%$ and $p = 9.56e{-}15$ for transparent overlay stimulus).

For C-shapes, Zhou et al. (2000) found 4 of 16 cells to be tuned significantly and consistent, 12 of 16 to be not significantly tuned, and no cells to be significantly tuned and inconsistent. Qiu and von der Heydt (2005) found 35% of 174 neurons in V2 to be selective to the side-of-figure of a luminance square, 40% to be selective to depth order, and 21% selective to both, of which 81% were consistent between luminance-defined and disparity-defined side-of-figure. For a transparent overlay stimulus, Qiu and von der Heydt (2007) found 127 of 244 of cells to be tuned to the side-of-figure of the luminance square and 30 of those cells to be significantly tuned to the side-of-figure of the transparent bars and consistent with the preferred side-of-figure for the luminance square; they did not report the number of significantly tuned, inconsistent cells.

Similar to the findings of Zhou et al. (2000), we found that a little more than half of all neurons were significantly selective to the side-of-figure of the single luminance squares (two-way ANOVA, $p < 0.01$). Compared with Zhou et al. (2000), however, we found a substantially higher proportion of cells that showed inconsistent border-ownership selectivity when comparing tuning for side-of-figure in luminance-defined figures and another artificial stimulus (Table 1). In particular, for C-shapes and overlapping outlines, we found even more cells that significantly preferred a side inconsistent with the single luminance square preference than cells that preferred a consistent side. For the overlapping luminance squares, the transparent stimulus, and its four square control, in turn, there were more cells that were consistently significantly selective than cells significantly preferring the opposite side, which is more consistent with the findings of

Zhou et al. (2000). Demonstrating that we were recording from the same general class of cells as von der Heydt's group, we reconstructed their population analysis for the transparent stimulus from Qiu and von der Heydt (2007) and found qualitatively similar results (Fig. 3): A subtle difference in the stimuli (Fig. 3, top) leads to the perception of either two transparent overlaid bars or four rounded squares, which changes the border ownership at the receptive field location. Although the scale of the marginal distributions of border-ownership modulation indices was somewhat different between our sample of cells and Qiu and von der Heydt (2007), we replicated the switching in border-ownership signaled by the cell population.

We next quantified how many cells were significantly tuned to a given number of artificial stimuli and consistent with the preferred side-of-figure for the single luminance square (Fig. 4a). We found that 40% of cells were not significantly and consistently tuned for any of the artificial stimuli, and not a single cell was significantly and consistently tuned for all artificial stimuli. There were 11 cells that were tuned consistently across all artificial stimuli, but not significant for all (i.e., the sign of the modulation index) averaged across the two contrast conditions, was the same for each artificial stimulus as for the single luminance square (standard test). In sum, we failed to find a population of cells that were significantly tuned to the side-of-figure across the whole large battery of artificial stimuli in a consistent manner; in particular, many ostensible border-ownership cells that were significantly tuned to the side-of-figure of luminance squares turned out to be inconsistent when tested with a battery of other stimuli. These results strongly suggest that border ownership is encoded
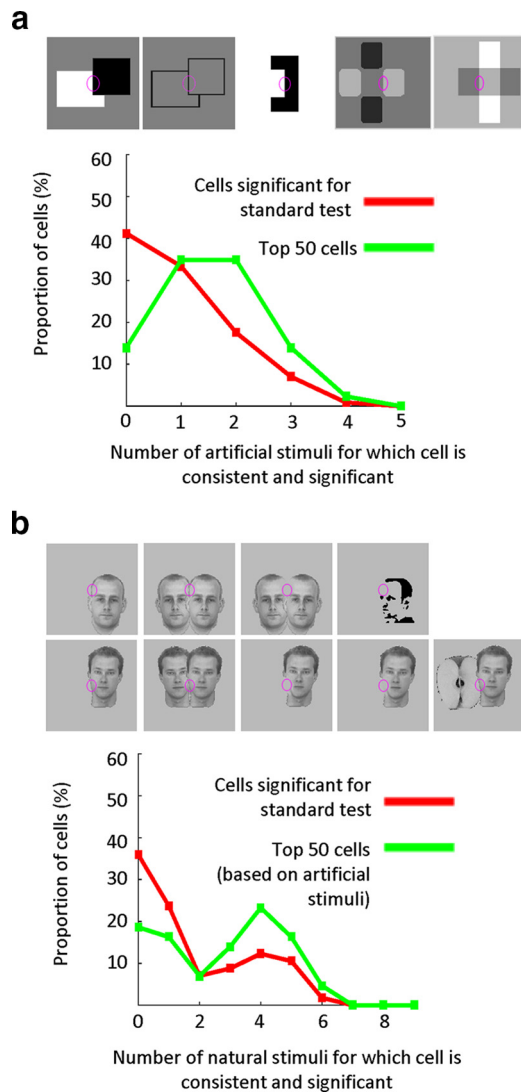
Figure 4. Quantification of response consistency across artificial and natural stimuli. Histograms represent the proportions of cells that are consistent and significant for a given number of artificial (**a**) and natural stimuli (**b**). The five artificial stimuli and nine natural stimuli based on which the consistency of the cell was tested are shown at the top of **a** and **b**, respectively. A cell is considered consistent for a given stimulus if the modulation index for that stimulus has the same sign as for the single luminance square (standard test). Significance criterion is $p < 0.01$. Red lines indicate the proportion of cells among all cells that were significantly tuned to the side-of-figure of the single luminance square. Blue lines only take into account the 50 cells with the highest modulation index across artificial stimuli.

by a population of cells, which are each, individually, imperfect in their border-ownership selectivity.

To determine whether, despite the inconsistencies of many cells, we could still extract a population signal that is consistent across all artificial stimuli, we next computed the average population response across the most consistent cells. We evaluated each neuron's consistency by computing the average modulation index across all artificial stimuli (shown in Figure 5a–d). The average modulation index across artificial stimuli had a mean of 0.038 and an SD of 0.081 across all 201 cells (see also histograms in Fig. 6, top). We then chose the 50 most consistent cells based on this average modulation index. The exact number of chosen cells did not qualitatively influence the results. Cells were pooled across V2 and V3; when comparing consistency to artificial stimuli between V2d and V3d, we found slightly but significantly more

consistent border-ownership signals in V3d than in V2d ($p = 0.02$, unpaired t test on average modulation indices across artificial stimuli based on 137 cells recorded from V2d and 64 cells recorded from V3d). As can be seen from Figure 4a (green line), by selecting the 50 top cells with this method, the proportion of cells that were consistent across a high number of artificial stimuli substantially increased. Figure 5a–d shows the average PSTHs of these 50 selected cells to artificial stimuli. By focusing on the most consistent border-ownership cells, we obtained a border-ownership signal that was reliable across artificial stimuli. The response difference for the standard luminance square test (Fig. 5a) was strong and consistent, which is expected because we defined the preferred side of each cell based on the luminance square responses. The responses to the other artificial stimuli (Fig. 5b–d) were also significantly stronger when presented on the preferred side, which is a sanity check that the population of 50 cells indeed represented consistent border-ownership cells. The population border-ownership signal for occluding outlines and C-shapes was weaker than for other artificial stimuli but was nevertheless consistent (see also Table 1).

Next, we asked how this population of border-ownership cells that responded consistently to the side-of-figure of artificial stimuli responded to a variety of natural stimuli. We found that the side-of-figure selectivity to artificial stimuli generalized to natural stimuli (Fig. 5e–k): When presenting the edge of a single face in the receptive field, the average response of border-ownership cells was significantly higher when the face was on the preferred side compared with when it was presented mirror-symmetrically on the nonpreferred side. Interestingly, the consistency of responses was even stronger and more significant than for most artificial stimuli based on which the cells were selected for. Moreover, the response was higher when the foreground face of two overlapping faces was presented on the preferred side. We also presented the isolated local part of the overlapping face stimulus to determine how much border-ownership information was given in the local edge alone. For the first pair of overlapping faces (Fig. 5f, top), we had chosen a central edge that was located on the forehead and fairly straight (Fig. 5g for the isolated stimulus and Fig. 5h for an enlarged version). For this pair, we could find no significant response difference for the isolated local stimulus alone. For the second pair, we chose a more convex edge at the chin and the local stimulus exhibited a T-junction at the bottom of the stimulus (Fig. 5f–h, bottom). This local stimulus by contrast did evoke a significant response difference that was consistent with the border-ownership of the cells. We next presented faces with illusory contours, namely, Mooney faces and the natural faces mentioned above with the local edge over the receptive field deleted. For each of these stimuli, we ensured that within the receptive field the stimulus was identical to the background presented during the 150 ms OFF time. As can be seen from Figure 5i–j, the response amplitude was much weaker when deleting the local stimulus, but the response difference was still significant and consistent with the border-ownership selectivity. When presenting a stimulus with the local edge not completely deleted but strongly reduced in contrast (Fig. 5j, bottom), the amplitude was in between the amplitudes to the full face and the locally deleted face, and again the response difference was consistent with the border ownership. We tested whether the response difference could be evoked by presenting the mere outlines of the overlapping faces (Fig. 5k, top) and found that the outlines alone did not evoke a significant response difference. The border-ownership selectivity for natural stimuli was not limited to faces, as a face occluding an apple also evoked consistent border ownership (Fig.
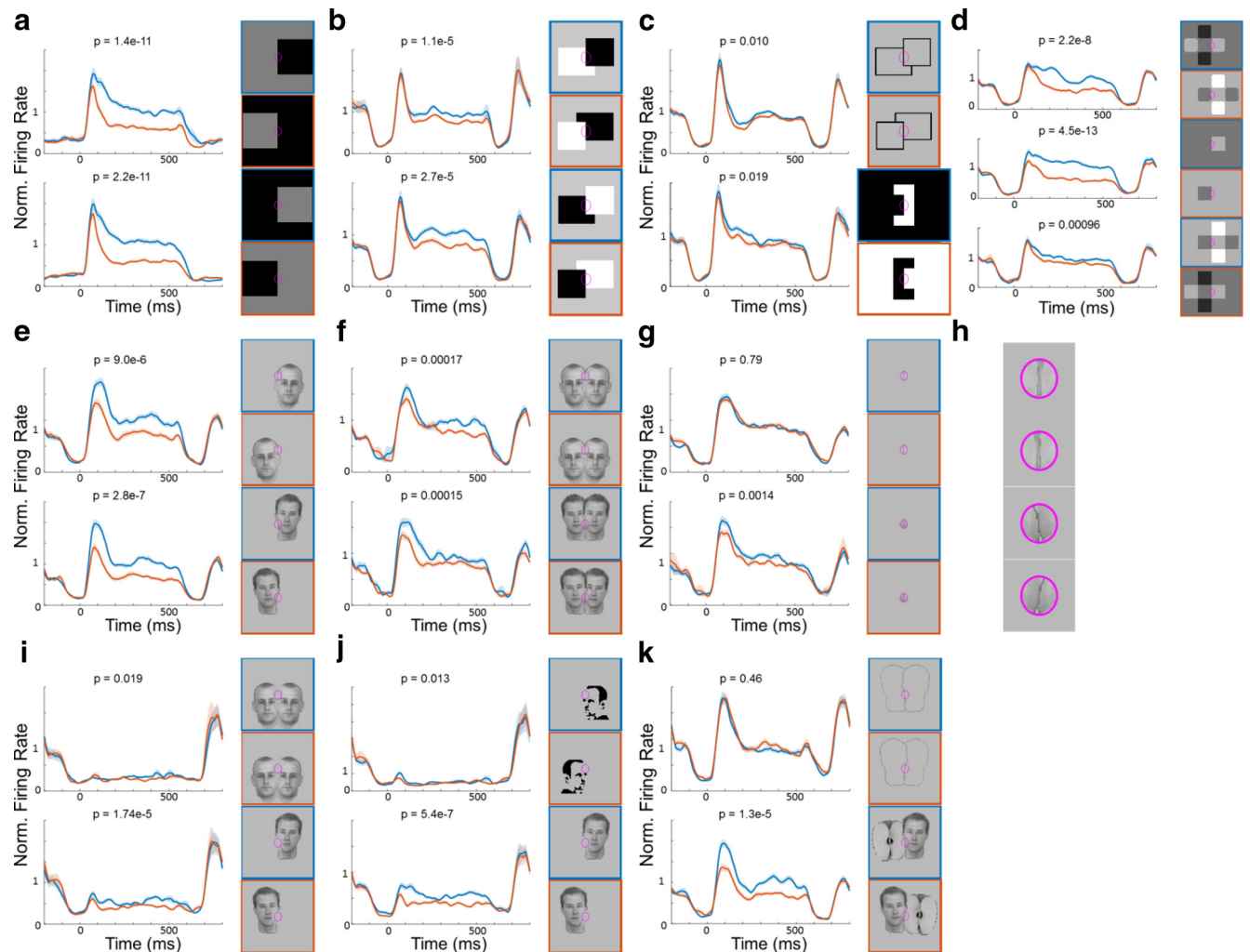
Figure 5. Responses of border-ownership cells to simple and natural stimuli. Left, The PSTHs represent the population mean responses (±1 SEM, shaded region) of 50 selected border-ownership cells to stimuli shown on the right outlined with the corresponding color. Both artificial and natural stimuli were rotated and positioned such that the central edge (indicated by purple ellipse) was in the preferred orientation and centered on the receptive field of each cell. Preferred side of figure of border-ownership cells was determined based on the response to simple luminance squares (shown in **a**) alone. Blue represents conditions where the figure was on the preferred side. Red represents matched conditions where the stimulus looks identical/similar in the receptive field (purple ellipse) but the figure is on the nonpreferred side. The 50 cells were determined to be consistent border-ownership cells based on their responses to simple, artificial stimuli alone. *p* values were determined using two-tailed paired *t* tests. **a**, Population mean responses to luminance squares. A higher response to the preferred side is expected as the preferred side was determined based on the response to the luminance square. **b**, Population mean responses to occlusion stimuli. **c**, Population mean responses to occluding outlines and C-shapes. **d**, Population mean responses to four squares (top), single squares (middle), and transparent overlay (bottom). **e**, Population mean responses to single faces. **f**, Population mean responses to overlapping faces. **g**, Responses to isolated edge stimulus of faces alone. This stimulus was generated by removing the part outside the receptive field for stimuli in **f**. **h**, Magnified version of stimuli in **g**. **i**, Population mean responses to occluding and single faces with illusory contour, where the local stimulus inside the receptive field was removed. **j**, Population mean responses to Mooney faces (top) and to stimulus on bottom of **f** but with low-contrast border, where the contrast was reduced locally inside the receptive field (bottom). **k**, Population mean responses to outlines of occluding faces and faces occluding an apple.

5*k*, bottom). However, for this stimulus, we cannot exclude that the response difference was due to the contrast difference between the two different objects. Figure 4*b* shows the proportions of cells that were consistent and significantly tuned to a given number of natural stimuli. Interestingly, among the 50 cells that were selected based on consistency for artificial stimuli (green line), there was a higher proportion of cells that was consistent and significant across many natural stimuli. This shows that the consistency across artificial stimuli generalizes to natural stimuli. For the natural stimuli in Table 1, the percentage of inconsistent cells substantially decreases when taking into account only the 50 cells with highest modulation indices averaged across artificial stimuli.

We next further examined the relationship between side-of-figure preference for artificial stimuli and natural stimuli (Fig. 6). In general, the modulation indices for artificial and natural stim-

uli showed small but strongly significant correlations, indicating that cells signaling the side-of-figure of artificial stimuli correctly also tend to correctly signal the side-of-figure of natural stimuli. Yet, there is considerable residual variance of natural stimuli modulation indices left, indicating that not every cell is consistent for every stimulus, but the correct side-of-figure needs to be inferred from the population activity. Both the luminance square modulation index and average modulation index across artificial stimuli were positively correlated with natural stimulus modulation indices; however, for both single and overlapping full faces, the average modulation index across artificial stimuli explained almost twice as much variance as the modulation index for luminance squares alone. Indeed, we found that, for most natural stimuli, the average modulation index across artificial stimuli explained more variance than the modulation indices of each individual artificial stimulus. This suggests that, for these natural
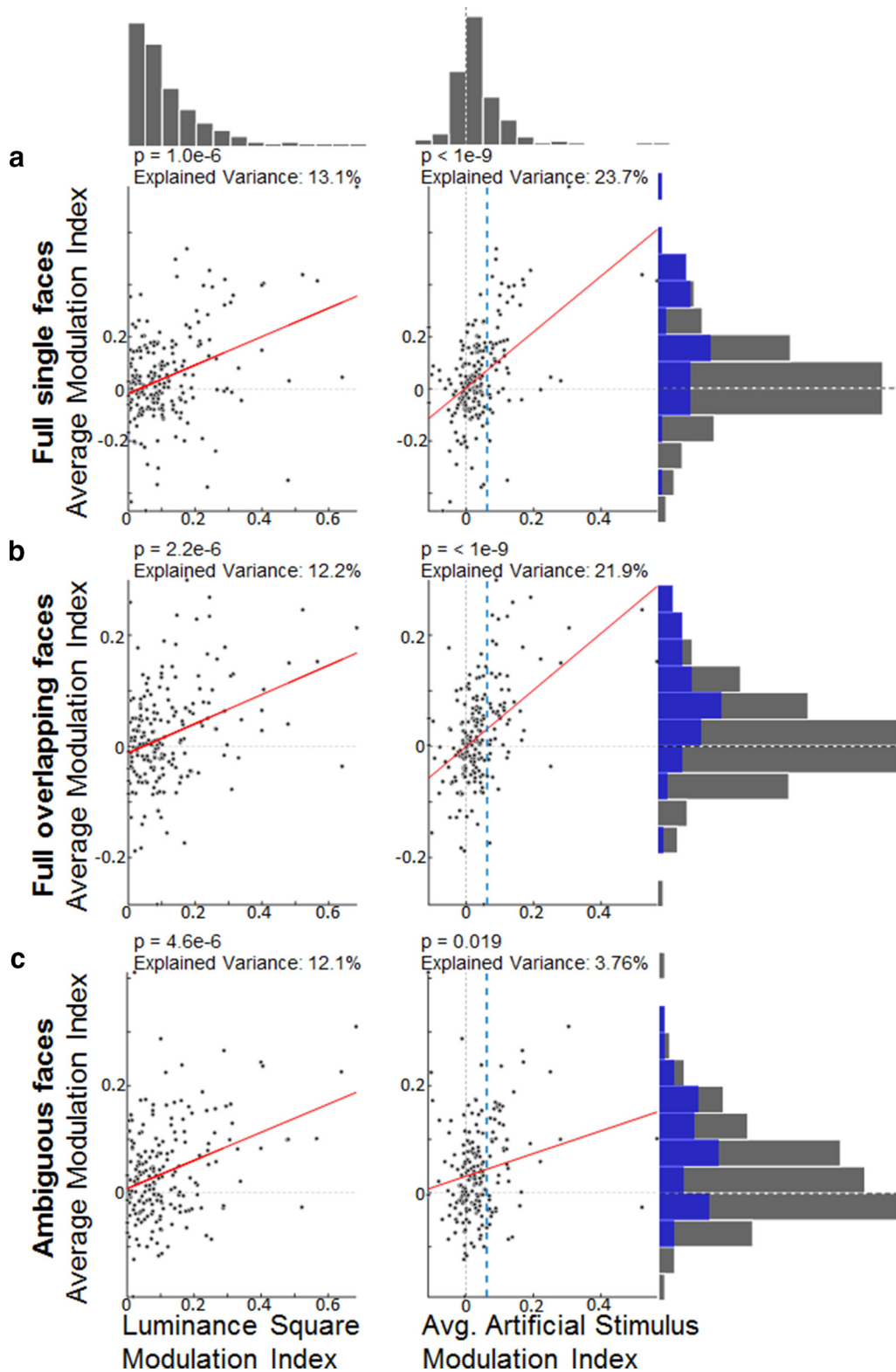
Figure 6. Correlations between modulation indices of artificial stimuli and modulation indices of natural stimuli. The six scatter plots show how responses to the side-of-figure of artificial stimuli correlate with responses to the side-of-figure of natural stimuli. Left, Scatter plots have the modulation indices of the single luminance square on the x-axis. Right, Scatter plots have the average modulation index across all artificial stimuli on the x-axis. Both types of modulation indices are plotted against the average modulation index across a class of natural stimuli on the y-axis. *a*, The first category includes full single faces shown in Figure 5e. *b*, The second category contains overlapping faces (Fig. 5f). *c*, The third category includes faces with ambiguous contours (Fig. 5i, j). Red lines indicate Standard Model I regression lines (i.e., the ordinary least-squares fit of the abscissa to the ordinate). p values were computed by comparing a transform of the Pearson correlation coefficient with a Student's t distribution. Using a permutation test on the Pearson correlation coefficient yielded qualitatively the same results. Vertical dashed blue line in the right plots indicates the average modulation index, which was the threshold for being chosen for the 50 selected cells used in Figure 5. Top, Right, Histograms represent the marginal distributions of average modulation indices for luminance squares, artificial stimuli, and natural stimuli, respectively, across all 201 cells (blue histograms show modulation indices for the 50 selected cells).
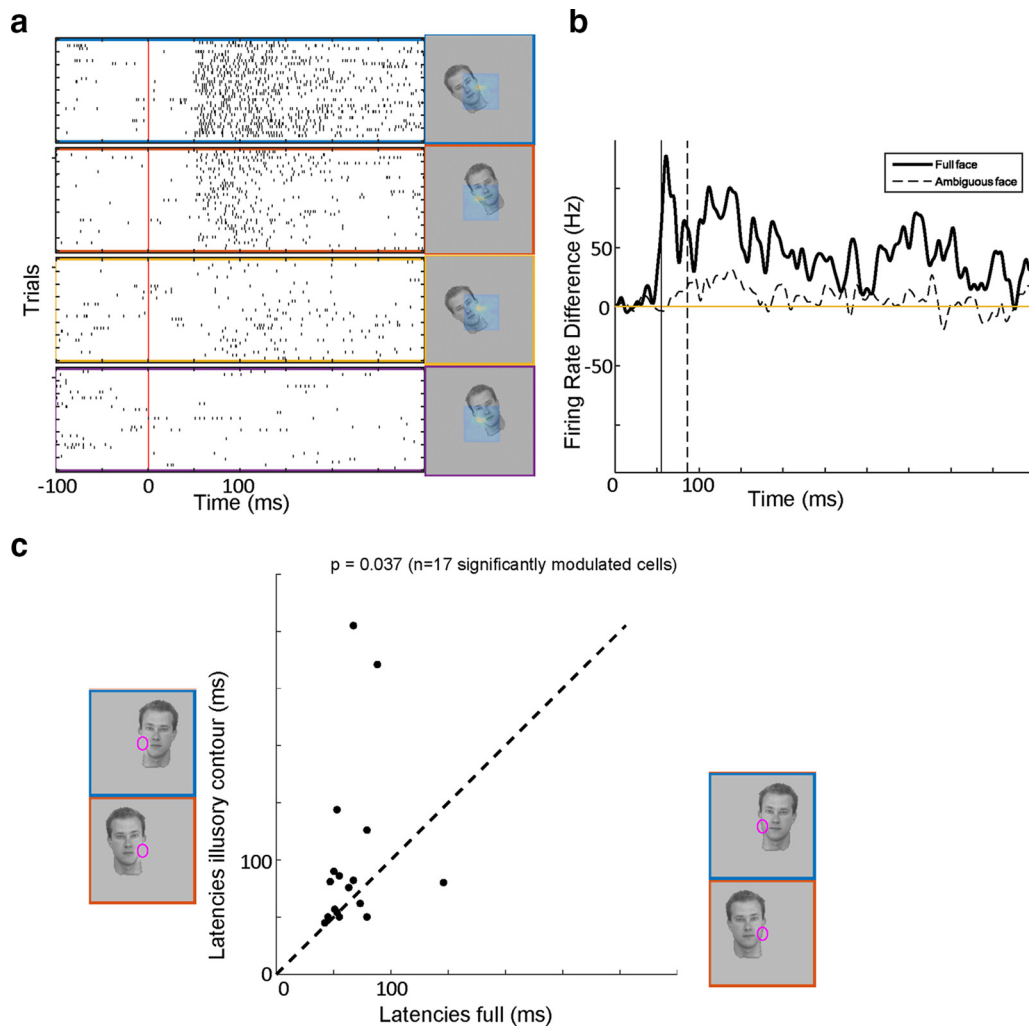
Figure 7. Latency differences between full stimuli and stimuli with illusory contours. *a*, Raster plots of example cell responses to full faces presented on its preferred side (blue) and nonpreferred side (red), respectively, and responses to the stimulus with locally deleted edge (yellow and purple). The preferred side was defined by the cell's responses to luminance squares. *b*, Difference of the PSTHs for preferred side and nonpreferred side, for the full face (solid line) and the face with locally deleted edge (dashed line). Vertical solid and dashed lines indicate the latencies for full and ambiguous faces, respectively, defined as the time the difference reaches half its peak value. *c*, Population scatter plot of latencies of border-ownership signal of full face stimulus shown at the *x*-axis against border-ownership latency for the stimulus with locally deleted edge shown at the *y*-axis. *p* value was computed with a paired Wilcoxon rank sum test.

stimuli, using a battery of artificial stimuli is a better predictor for whether a cell is a border-ownership cell and will respond consistently to these natural stimuli than just using the luminance square standard test alone. In contrast, for the stimuli with ambiguous contours, the average modulation index across artificial stimuli was a worse predictor than the modulation index for the standard test alone. This could suggest that the populations of border-ownership cells that represent the side-of-figure of illusory contours might not always carry complete information about the correct segmentation of certain artificial stimulus conditions. However, for the single ambiguous face (Fig. 5*i*, bottom), the modulation index averaged across artificial stimuli actually explained more variance than the single luminance square modulation index. It is noteworthy that the single ambiguous face was the stimulus for which none of the top 50 cells was inconsistent (Table 1), whereas for artificial stimuli there tended to be many inconsistent cells. For other natural stimuli, the number of inconsistent cells was also low compared with artificial stimuli. There were 3 of 50 cells that were significantly tuned to the single full face stimulus but inconsistent with the luminance square stimulus and no cells that were significantly tuned to

the single full face stimulus and inconsistent with the single ambiguous face.

Justifying our choice to focus on the most consistent cells, the consistency between border-ownership selectivity for artificial and natural stimuli was greatly weakened when we analyzed all cells regardless of their consistency across artificial stimuli (compare marginal distributions using all cells [gray bars] with the distributions using only the top 50 cells [blue bars] in Figure 6). When we included all cells, the population average could not significantly determine the border-ownership of occluding faces. For single faces, there was still a consistent border-ownership signal left, even when averaging across all cells, but it was less significant than when averaging across the top cells (despite larger sample size).

We compared latencies of the border-ownership signal for the full face stimulus and the ambiguous face stimulus. Figure 7*a* shows raster plots of an example cell's responses to full and ambiguous face stimuli. As can be seen from the difference of the PSTHs for the preferred and nonpreferred condition, the border-ownership signal for the stimulus with illusory contour is smaller and delayed. As discussed in the previous paragraph, the 50 cells

most consistent for artificial stimuli formed a population that signaled border-ownership of the ambiguous stimulus reliably. Thus, we compared latencies for the single full face stimulus and corresponding single ambiguous stimulus for the 17 of 50 cells that were consistently and significantly border-ownership selective for both stimuli. The latency for the full stimulus was on average 65 ms and significantly shorter than the latency to the stimulus with illusory contour, which was on average 100 ms.

Discussion

We were interested in how border-ownership cells, which have been shown to respond consistently to the side-of-figure of artificial stimuli, respond to natural stimuli and stimuli with ambiguous contours.

In our study, we probed responses of a large number of V2 and V3 cells for border-ownership selectivity across a large battery of different stimuli, both artificial and natural, enabling us to rigorously assess the extent to which each cell showed consistent tuning across different stimulus conditions. As can be seen from Table 1, we found many cells that were ostensibly border-ownership cells based on selectivity to the side-of-figure of single luminance squares but responded inconsistently to other artificial stimuli. Indeed, there was not a single border-ownership cell (of 201 tested) that was significantly tuned and consistent to every single artificial stimulus. The considerable number of partially inconsistent cells in Table 1 contradicts the simplistic concept of the perfect border-ownership cell, despite the intuitive appeal of a single cell explaining a variety of perceptual phenomena. Instead, it seems that single border-ownership cells carry incomplete information about figure-ground segmentation for only a subset of conditions. Thus, in order for the brain to reliably determine the correct segmentation of the scene, it needs to average the activity of multiple border-ownership cells that each carry information about border-ownership in different situations. Indeed, we found that, by averaging across the most consistent cells, it was possible to get a population signal that was reliable across artificial stimuli.

We found that this population signal also consistently signaled the side-of-figure of a battery of natural face and object stimuli. This further supports the notion that border-ownership cells play a vital role for segmenting the visual scene into objects. For rather simple stimuli, such as the single faces, one could argue that the consistent response might simply be caused by asymmetric receptive fields such that the cells prefer more complex texture on one side of their receptive field center. In contrast, the consistent response to overlapping faces, which are visually very similar in both conditions, indicates that border-ownership cells are indeed inferring the side of the foreground object. For the pair of overlapping faces at the bottom of Figure 5g, where the local edge is convex and contains a telling T-junction, even the isolated local stimulus evoked a consistent, though smaller, response difference (Fig. 5h, bottom). This is consistent with psychophysical and computational evidence by Fowlkes et al. (2007) that the local bottom-up cues of borders in natural scenes are in many situations enough to decide border ownership. On the other hand, when the local stimulus is ambiguous and does not evoke a significant response difference (Fig. 5h, top), global context cues can help to decide border ownership (Fig. 5g, top). To test whether shape of the object is sufficient for border-ownership cells to determine the side-of-figure, we presented the outlines of the overlapping faces alone. This did not evoke a significant response difference, which is consistent with the subjective experience that the border ownership of the outlines alone is ambiguous without the texture and other features of the face. These results suggest

that border-ownership cells integrate a variety of the object's features, including local cues, as well as shape and texture outside the receptive field.

Illusory contours have been a major subject of study for figure-ground segmentation in psychological literature (Heitger et al., 1994), and von der Heydt et al. (1984) found neurons in V2 that responded to illusory contours; however, the amplitude was smaller and the latency 10 ms longer compared with real contours (von der Heydt and Peterhans, 1989). The existence of such cells in V2 and also V1 was confirmed by several other studies (von der Heydt and Peterhans, 1989; Grosof et al., 1993; Sheth et al., 1996; Lee and Nguyen, 2001; Ramsden et al., 2001), and Bakin et al. (2000) found that they also respond to depth-defined illusory contours. Yet, up to now, it has not been known whether border-ownership cells consistently signal the side-of-figure for illusory contours (Kogo and Wagemans, 2013), although there have been several computational models (Finkel and Sajda, 1992; Sajda and Finkel, 1992; Kogo et al., 2010) where border-ownership and illusory contours both emerge from a dynamic network computing figure-ground organization. Recently, there has also been a discussion paper (Kogo and Wagemans, 2013), suggesting the intertwinedness of illusory contours with border ownership, which received many commentaries. Among them, von der Heydt (2013) argued that illusory contours and border ownership might be represented by distinct populations in V2. Here, we have shown that border-ownership cells do consistently signal the side-of-figure for illusory contours, even though local edge information is missing. The number of cells significantly signaling the correct side-of-figure for illusory contours was slightly lower than for full contours (31 of 50 significant cells for full stimulus vs 20 of 50 significant cells for ambiguous stimulus, unpaired t test, $p < 0.05$), which may be partly due to the reduced amplitude of responses to illusory contours in general. Among the 50 cells most consistent for artificial stimuli, there were only 3 cells that were significantly tuned for the full face stimulus but inconsistent for the luminance square, and no cells that were significantly and inconsistently tuned to the side-of-figure of the ambiguous stimulus and the luminance square. It is surprising that the ambiguous face stimulus turned out to be the stimulus where no cell signaled inconsistent border ownership. A possible explanation is that the local stimulus in the classical receptive field causes the transient response, which can lead to errors for the unambiguous stimuli, whereas for the ambiguous face stimulus the border-ownership signal is evoked entirely by feedback from context (Gilbert and Li, 2013), which is more reliable. Analogously to the onset responses of illusory contours studied by von der Heydt and Peterhans (1989), the response differences signaling border ownership of illusory contours were also smaller and delayed by ~30 ms compared with the full stimuli.

Previously, the responses of border-ownership cells have been modeled as a result of pure feedforward operations (Heitger et al., 1994; Sakai and Nishimura, 2006; Supèr et al., 2010), of intraareal dynamics (Baek and Sajda, 2005; Zhaoping, 2005) and of feedforward and feedback interactions between areas (Craft et al., 2007; Jehee et al., 2007; Kogo et al., 2010), respectively. Our findings falsify pure feedforward models as they predict the same latencies of border-ownership signals regardless of the stimulus. Instead, recurrent (Lamme and Roelfsema, 2000) connections might need to be used to resolve ambiguous scenes. It is conceivable that the latency of the border-ownership signal is increased because intra-areal network dynamics or feedforward-feedback interactions between V2 and V4 require more iterations to resolve the border ownership. However, at least for the Mooney

face, which evokes a significant and consistent border-ownership signal (Fig. 5*j*), it seems unlikely that contour completion mechanisms building from low-level cues would be sufficient to infer the illusory boundary of the face, but instead knowledge about face shape appears to be required. Assuming that the border-ownership signal that emerges ~100 ms for the full face originates from feedback from V4, it is also possible that the later signal for the border-ownership of the illusory contour arises from feedback from an area later in the hierarchy (e.g., posterior IT). Such corticocortical feedback loop interactions have been shown to exist between V1 and V4 (Chen et al., 2014). The most posterior face patch PL has latencies of 80 ms to distinguish faces from objects (Issa and DiCarlo, 2012) and is thus a possible candidate feedback source.

Previously, there has been a debate (Vecera and Farah, 1997; Vecera and O'Reilly, 1998; Peterson, 1999) on whether segmentation precedes recognition or vice versa. Our results (Fig. 7) suggest the possibility of a third alternative that is consistent with models of Bayesian inference (Rao and Ballard, 1999; Lee and Mumford, 2003; Yuille and Kersten, 2006): According to this hypothesis, initially, retinotopic areas try to segment the scene into regions corresponding to objects based on low-level cues, potentially using a combination of feedforward mechanisms using local cues as T-junctions, intra-areal and interareal dynamics. However, because segmentation of natural scenes is inherently ambiguous based on low-level cues and not every edge is an object border, this initial segmentation wave can only make guesses about which regions correspond to objects and relay these regions to IT. IT then attempts to recognize objects in the hypothesized regions and can accept or falsify the hypotheses by exciting or inhibiting border-ownership cells via feedback. In this way, IT would generate a representation of object surfaces in retinotopic cortex. However, the found latencies are merely suggestive evidence, and simultaneous recordings and perturbations of multiple areas will be necessary to dissect the exact mechanisms.

Overall, we found that many ostensible border-ownership cells, as determined by the single luminance square, turned out to be inconsistent for one or more stimuli when presented with a larger battery of artificial stimuli, and not a single cell showed consistent border-ownership preference across all stimulus conditions tested. This emphasizes the necessity for future studies to present a larger set of stimuli to identify the most consistent border-ownership cells and the need to use a population code for decoding segmentation. Importantly, the population of border-ownership cells that was consistent across most artificial stimuli could also reliably segment both natural face and object stimuli and, with some delay, even ambiguous stimuli where local edge information was completely missing, which suggests that border-ownership cells integrate both local, low-level cues and global, high-level object cues to segment the visual scene. By exploiting new techniques, such as population calcium imaging, optogenetics, and simultaneous recordings in retinotopic and IT cortex, future work might be able to reveal how border ownership cell populations are read out by downstream areas across different stimulus configurations.

References

Baek K, Sajda P (2005) Inferring figure-ground using a recurrent integrate-and-fire neural circuit. IEEE Trans Neural Syst Rehabil Eng 13:125–130. CrossRef Medline

Bakin JS, Nakayama K, Gilbert CD (2000) Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. J Neurosci 20:8188–8198. Medline

Chen M, Yan Y, Gong X, Gilbert CD, Liang H, Li W (2014) Incremental integration of global contours through interplay between visual cortical areas. Neuron 82:682–694. CrossRef Medline

Craft E, Schütze H, Niebur E, von der Heydt R (2007) A neural model of figure-ground organization. J Neurophysiol 97:4310–4326. CrossRef Medline

Driver J, Baylis GC (1996) Edge-assignment and figure-ground segmentation in short-term visual matching. Cogn Psychol 31:248–306. CrossRef Medline

Finkel LH, Sajda P (1992) Object discrimination based on depth-from-occlusion. Neural Comput 4:901–921. CrossRef

Fowlkes CC, Martin DR, Malik J (2007) Local figure-ground cues are valid for natural images. J Vis 7:2. CrossRef Medline

Freiwald WA, Tsao DY (2010) Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science 330:845–851. CrossRef Medline

Gilbert CD, Li W (2013) Top-down influences on visual processing. Nat Rev Neurosci 14:350–363. CrossRef Medline

Grill-Spector K, Kanwisher N (2005) Visual recognition as soon as you know it is there, you know what it is. Psychol Sci 16:152–160. CrossRef Medline

Grosof DH, Shapley RM, Hawken MJ (1993) Macaque V1 neurons can signal illusory contours. Nature 365:550–552. CrossRef Medline

Heitger F, von der Heydt R, Kübler O (1994) A computational model of neural contour processing: figure-ground segregation and illusory contours. From perception to action conference, pp 181–192. Lausanne, Switzerland: IEEE.

Issa EB, DiCarlo JJ (2012) Precedence of the eye region in neural processing of faces. J Neurosci 32:16666–16682. CrossRef Medline

Jehee JF, Lamme VA, Roelfsema PR (2007) Boundary assignment in a recurrent network architecture. Vision Res 47:1153–1165. CrossRef Medline

Kiani R, Esteky H, Tanaka K (2005) Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. J Neurophysiol 94:1587–1596. CrossRef Medline

Kogo N, Wagemans J (2013) The "side" matters: how configurality is reflected in completion. Cogn Neurosci 4:31–45. CrossRef Medline

Kogo N, Strecha C, Van Gool L, Wagemans J (2010) Surface construction by a 2-D differentiation-integration process: a neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures. Psychol Rev 117:406–439. CrossRef Medline

Lamme VA, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci 23:571–579. CrossRef Medline

Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex. J Opt Soc Am A Opt Image Sci Vis 20:1434–1448. CrossRef Medline

Lee TS, Nguyen M (2001) Dynamics of subjective contour formation in the early visual cortex. Proc Natl Acad Sci U S A 98:1907–1911. CrossRef Medline

Marr D (1982) Vision: a computational investigation into the human representation and processing of visual information. New York: Henry Holt.

Maunsell JH, Gibson JR (1992) Visual response latencies in striate cortex of the macaque monkey. J Neurophysiol 68:1332–1344. Medline

McDermott J (2004) Psychophysics with junctions in real images. Perception 33:1101–1127. CrossRef Medline

Nakayama K, He ZJ, Shimojo S (1995) Visual surface representation: a critical link between lower-level and higher-level vision. Vis Cogn 2:1–70. CrossRef

Ohayon S, Tsao DY (2012) MR-guided stereotactic navigation. J Neurosci Methods 204:389–397. CrossRef Medline

Pack CC, Born RT, Livingstone MS (2003) Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. Neuron 37:525–535. CrossRef Medline

Peterson MA (1999) What's in a stage name? Comment on Vecera and O'Reilly (1998). J Exp Psychol 25:276–286.

Peterson MA, Gibson BS (1993) Shape recognition inputs to figure-ground organization in three-dimensional displays. Cogn Psychol 25:383–429. CrossRef

Peterson MA, Gibson BS (1994) Must figure-ground organization precede object recognition? An assumption in peril. Psychol Sci 5:253–259. CrossRef

Peterson MA, Kim JH (2001) On what is bound in figures and grounds. Visual Cogn 8:329–348. CrossRef

Qiu FT, von der Heydt R (2005) Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. Neuron 47:155–166. CrossRef Medline

Qiu FT, von der Heydt R (2007) Neural representation of transparent overlay. Nat Neurosci 10:283–284. CrossRef Medline

Ramsden BM, Hung CP, Roe AW (2001) Real and illusory contour processing in area V1 of the primate: a cortical balancing act. Cereb Cortex 11:648–665. CrossRef Medline

Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87. CrossRef Medline

Riesenhuber M, Poggio T (2002) Neural mechanisms of object recognition. Curr Opin Neurobiol 12:162–168. CrossRef Medline

Rubin E (1958) Figure and ground. Readings Percept 194–203.

Sajda P, Finkel LH (1992) Simulating biological vision with hybrid neural networks. Simulation 59:47–55. CrossRef

Sakai K, Nishimura H (2006) Surrounding suppression and facilitation in the determination of border ownership. J Cogn Neurosci 18:562–579. CrossRef Medline

Sheth BR, Sharma J, Rao SC, Sur M (1996) Orientation maps of subjective contours in visual cortex. Science 274:2110–2115. CrossRef Medline

Sugihara T, Qiu FT, von der Heydt R (2011) The speed of context integration in the visual cortex. J Neurophysiol 106:374–385. CrossRef Medline

Supèr H, Romeo A, Keil M (2010) Feed-forward segmentation of figure-ground and assignment of border-ownership. PLoS One 5:e10705. CrossRef Medline

Tsao DY, Vanduffel W, Sasaki Y, Fize D, Knutsen TA, Mandeville JB, Wald LL, Dale AM, Rosen BR, Van Essen DC, Livingstone MS, Orban GA, Tootell RB (2003) Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. Neuron 39:555–568. CrossRef Medline

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. Science 311:670–674. CrossRef Medline

Tsao DY, Moeller S, Freiwald WA (2008) Comparing face patch systems in macaques and humans. Proc Natl Acad Sci U S A 105:19514–19519. CrossRef Medline

Vecera SP, Farah MJ (1997) Is visual image segmentation a bottom-up or an interactive process? Percept Psychophys 59:1280–1296. CrossRef Medline

Vecera SP, O'Reilly RC (1998) Figure-ground organization and object recognition processes: an interactive account. J Exp Psychol Hum Percept Perform 24:441–462. Medline

von der Heydt R (2013) Neurophysiological constraints on models of illusory contours. Cogn Neurosci 4:49–50. CrossRef Medline

von der Heydt R, Peterhans E (1989) Mechanisms of contour perception in monkey visual cortex: I. Lines of pattern discontinuity. J Neurosci 9:1731–1748. Medline

von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. Science 224:1260–1262. CrossRef Medline

Yuille A, Kersten D (2006) Vision as Bayesian inference: analysis by synthesis? Trends Cogn Sci 10:301–308. CrossRef Medline

Zhaoping L (2005) Border ownership from intracortical interactions in visual area V2. Neuron 47:143–153. CrossRef Medline

Zhou H, Friedman HS, von der Heydt R (2000) Coding of border ownership in monkey visual cortex. J Neurosci 20:6594–6611. Medline