

Representation of object continuity in the visual cortex

Philip O'Herron

Krieger Mind/Brain Institute and Department of
Neuroscience, Johns Hopkins University,
Baltimore, MD, USA



Rüdiger von der Heydt

Krieger Mind/Brain Institute and Department of
Neuroscience, Johns Hopkins University,
Baltimore, MD, USA



An amazing feature of our visual system is the ability to detect and track objects in the stream of continually changing retinal images. Theories have proposed that the system creates temporary internal representations that persist across changing images, providing continuity. However, how such representations are formed in the brain is not known. Here we examined the time course of the responses of border-ownership-selective neurons in the visual cortex to displays that portray object continuity. We found that the neurons signal border ownership immediately when new objects appear, but when a border that has been assigned to one object is reassigned to another object while the first remains in the display, the initial responses persist. The neurons continue to signal the initial assignment despite the presence of contradicting figure-ground cues. We propose that border ownership selectivity reflects mechanisms that create object continuity.

Keywords: electrophysiology, perceptual organization, object recognition

Citation: O'Herron, P., & von der Heydt, R. (2011). Representation of object continuity in the visual cortex. *Journal of Vision*, 11(2):12, 1–9, <http://www.journalofvision.org/content/11/2/12>, doi:10.1167/11.2.12.

Introduction

Despite continual fluctuations of the retinal image, we perceive a stable world in which objects have continuity. We can identify objects across a sequence of changing images. We can find objects with the specific features we are looking for, and we can also distinguish and scrutinize objects we have not seen before. To explain these amazing abilities, theories assume that the brain creates temporary internal representations that link the elemental features of an object and persist across changing images to provide the necessary continuity (Kahneman, Treisman, & Gibbs, 1992; Pylyshyn & Storm, 1988; Rensink, 2000). However, how such representations are formed in the brain is not known.

Neurophysiological studies have shown that the early stages of the visual cortex produce maps of local features, but recent studies related to figure-ground organization have shown that there is also more global processing (e.g., Lamme, 1995; Lee, Mumford, Romero, & Lamme, 1998; Roelfsema, Lamme, & Spekreijse, 1998; Zhou, Friedman, & von der Heydt, 2000; Zipser, Lamme, & Schiller, 1996). While most neurons in areas V1 and V2 respond to local contrast borders and are orientation selective, about half of the neurons in V2 are also selective for the side on which a border is “owned” by a figure (border ownership, Zhou et al., 2000). The left-hand side of a square, for example, produces high firing rates in neurons of figure-right preference and low firing rates in neurons of figure-left

preference. Although these neurons can see only a small segment of border through their classical receptive field, they seem to “know” that this segment is part of the contour of a larger object. They integrate global shape information with various local cues, such as stereoscopic depth and occlusion cues, to infer which side is foreground and which side is background (Qiu & von der Heydt, 2005, 2007; von der Heydt, Qiu, & He, 2003; Zhang & von der Heydt, 2010; Zhou et al., 2000).

Figure-ground organization has two aspects. It shows that the system uses figure-ground cues to infer the ordering of objects in depth. A display of an L-shaped region next to a square region is perceived as one square overlapping another square; a texture region moving through a textured surround is perceived as a moving surface *in front of* a background surface. In these cases, the system uses figure-ground cues to infer depth, the dimension that is missing in images. However, figure-ground organization also gives insight into the way the system defines objects. The visual image is nothing but a large array of retinal cone signals that vary in time. Perception interprets this stream of signals as objects in space. A display can be as simple as a square region of pixels of one color surrounded by pixels of a different color, but perception interprets the square as an object, and the color boundaries as the contours of the object. Thus, the rules of figure-ground perception reveal something about how the system defines objects.

These two aspects are related, but there is an important difference. The determination of depth order involves the

evaluation of figure–ground cues, whereas object representation also requires short-term memory, because the system needs to represent object continuity. When a figure disappears and shortly after that another figure appears in a different location, we usually perceive one moving object. Thus, the second figure is not represented as a new object but merely as a different state of an existing object. Even in more complex displays, the system is able to keep track of objects (Pylyshyn & Storm, 1988).

We have previously found that border ownership signals persist when the edge of a square is replaced by an ambiguous edge (a split circular field; O'Herron & von der Heydt, 2009). However, from this study, it was not clear if the persistence reflects a slow decay of depth order signals or the persistence of an object representation. Depth order signals might decay slowly in the absence of new depth information. Alternatively, when the square is replaced with an ambiguous edge, a representation of the square might persist, and the edge, which coincides with one side of the square, might still be part of the persisting square representation. Our previous experiments also showed that the persisting border ownership signals at the ambiguous edge could be reversed quickly if a new figure was presented on the opposite side of the edge. Thus, the signals persisted in the absence of figure–ground cues but could be reset immediately by the presentation of new figure–ground information. This suggested that the absence of figure–ground cues is critical for the persistence. However, the experiment did not dissociate the effect of figure–ground cues from that of the hypothetical object representation because the display sequence did not portray object continuity. If such a representation was indeed formed and persisted into the ambiguous edge phase, it might have been overwritten when the new figure was presented.

In the experiments to be described, we pitted figure–ground cues against the hypothetical object representation. This was achieved by presenting two figures, one of which was moved so that the figure–ground cues reversed while the figures remained the same. Thus, the border between the figures was first assigned to figure A, and then to figure B, while both figures remained in the display. The question we asked was: will the responses to the new figure–ground cues be affected by the previous assignment to figure A? The results clearly show that this was the case. It took more than a second until the new figure–ground cues won the border over from the previous ownership.

Materials and methods

All animal procedures conformed to National Institutes of Health and USDA Guidelines as verified by the Animal Care and Use Committee of the Johns Hopkins University.

We studied neurons in two male adult rhesus monkeys (*Macaca mulatta*). The details of our general methods have been described (O'Herron & von der Heydt, 2009).

Preparation

The animals were prepared by implanting, under general anesthesia, first three small posts for head fixation, and later two recording chambers (one over each hemisphere). Fixation training was achieved by controlling fluid intake and using small amounts of juice or water to reward steady fixation.

Stimuli and experimental design

Stimuli were generated with Open Inventor on a Pentium 4 Linux workstation with NVIDIA GeForce 6800 graphics card using the anti-aliasing feature of the software and were presented on a 21-inch EIZO FlexScan T965 color monitor with 1600×1200 resolution at 72-Hz refresh rate. Stereoscopic pairs were presented side by side and superimposed optically at 40-cm viewing distance. The field of view subtended 17 by 26 deg visual angle. A white (93 cd/m^2) cross inside a 20-arcmin-diameter disk of 9 cd/m^2 served as fixation point. The color tuning of each neuron was determined with stationary flashing bars, and the minimum response field was mapped with bars and drifting gratings. Orientation and disparity tunings were determined with moving bars. The square figures were typically 4 deg on a side. Occasionally, larger figures were used, so that the figure was at least twice the linear size of the receptive field. The L-figure (Figure 1A) was created by adding a small square whose sides measured one-fourth the sides of the square. The figures had a texture of small random dots added to their surfaces in order to enhance the effect of motion on figure–ground assignment. The bottom figure moved at an angle 45 deg to the edge in the receptive field (the preferred orientation of the neuron). The distance of movement was $2/3$ of the length of the side of the square and the speed of movement was set by forcing the movement to last 0.5 s.

The direction of gaze was monitored for one eye with an infrared video-based system (Iscan ETL-200) at 60 Hz with a spatial resolution of 5120 (H) and 2560 (V). The eyes were imaged through an infrared-reflecting mirror, placing the camera on the axis of fixation. The optical magnification in our system resulted in a resolution of the corneal position signal of 0.08 deg visual angle in the horizontal and 0.16 deg in the vertical. Noise and drifts of the signal of course reduced the accuracy. Behavioral trials began with the presentation of the fixation mark on a blank screen. A test sequence was initiated when gaze was in a predetermined fixation window (1 deg radius) and the

first stimulus appeared 300 ms after fixation was detected. The monkey was rewarded for keeping its gaze in the fixation window for a fixed duration of 2.3 or 3.3 s, depending on the experiment. After successful termination of a trial, the display was blanked for an interval of 0.5 to 1.2 s. When fixation was broken, the trial was terminated and the following inter-trial interval was increased by 1 s.

Each of the experiments described involved variation of several stimulus parameters. For example, the moving figures experiment represented in [Figures 1A](#) and [1B](#) involved 4 binary variables: the local contrast polarity, the side of the receptive field on which the static figure was presented, whether the foot of the L was above or below the receptive field, and the two conditions (CUE REVERSAL or CONSISTENT, i.e., [Figure 1A](#) or [1B](#)). The ONSET condition ([Figure 1C](#)) was presented in a separate test and included two local contrast polarities and two directions of overlap. The FLIP condition ([Figure 1D](#)) was run in another test and included two contrast polarities and both sides of initial figure presentation. Factorial designs were used and all conditions of a test were presented in pseudo-random order in which each condition was presented once before moving on to the next repetition.

Recording procedures

Single-neuron activity was recorded extracellularly with epoxy-insulated tungsten microelectrodes inserted through the dura mater. A spike detection system (Alpha Omega MSD 3.22) was used. Spike times, stimulus events, and behavioral events were digitized and recorded by computer.

Cells in area V2 were recorded either in the lunate sulcus after passing through V1 and the white matter or in the lip of the post-lunate gyrus. The eccentricities of the receptive fields ranged from 0.74 to 6.8 deg (median 2.2 deg). After isolating a cell, we first characterized its selectivity for color, bar size, and orientation and mapped its receptive field using hand- and computer-controlled stimuli (Zhou et al., 2000). Next, border ownership selectivity was determined by a standard test using the edge of a square, square sizes of 3 and 8 deg, and both contrast polarities (Qiu & von der Heydt, 2005). If a cell was color selective, the preferred color and a 28 cd/m² gray were used for the two figure colors; otherwise, white (93 cd/m²) and gray (28 cd/m²) were used. The background color was the average of the figure colors, except in the FLIP condition ([Figure 1D](#)), in which the preferred color and the gray were used as the figure and background colors. The color of the blank screen shown between trials was also the average of the two colors. If a cell showed border ownership selectivity for single squares, it was then tested for selectivity to the overlapped condition and subjected to the tests of [Figure 1](#). In these tests, the figure edge length was set to be at least twice the size of the RF. Neurons that showed no border ownership selectivity in

the preliminary standard test (ANOVA, $p \geq 0.05$; about half of the neurons) were generally not tested further.

Data analysis

A total of 49 cells were tested in our experiments. Of these, 34 showed selectivity for border ownership in the CONSISTENT condition of [Figure 1B](#) (10 from monkey TH and 24 from monkey JA), as determined by a significant difference in the firing rates in the interval from 0 to 1.3 s after stimulus onset (ANOVA $p < 0.05$). Of these 34 neurons, 27 were also tested with the FLIP condition in [Figure 1D](#) (6 from monkey TH and 21 from monkey JA).

For the time course plots ([Figures 2B](#) and [4B](#)), we computed an average of the peristimulus time histograms (2-ms bin width) of the single neurons. Only border ownership selective cells ($p \leq 0.05$) were included in the average. The resulting averaged firing rates were smoothed with a Gaussian kernel of $\sigma = 20$ ms.

To describe the time course of the border ownership signal, we calculated fits to the population average (2-ms bin width) using multiphase least-squares approximation. For the ONSET condition, the fit had two phases: (1) a zero line and (2) a sum of two exponentials with independent time constants, amplitudes, and asymptotes. For CONSISTENT, CUE REVERSAL, and FLIP, the fit consisted of three phases: (1) a zero line, (2) a sum of two exponentials with independent time constants, amplitudes, and asymptotes, and (3) a third phase that differed between conditions. For the CUE REVERSAL and CONSISTENT conditions, the third phase were exponentials with individual time constants but the same asymptote, where the amplitudes were constrained so as to have continuity with the second phase. For FLIP, the third phase was a sum of two exponentials with independent time constants, amplitudes, and asymptotes, the amplitudes being constrained to achieve continuity with the second phase. The time points of transition between phases were additional free parameters. However, the transition times from phase 2 to phase 3 for conditions CONSISTENT and CUE REVERSAL (for which estimates had a large uncertainty) were assumed to be the same as that for the FLIP condition. The equations of the functions fitted to the data for CONSISTENT and CUE REVERSAL are given as follows:

$$y = \begin{cases} \left. \begin{aligned} &0 \cdot \text{phase one } (t) \\ &+ [c_1(\exp(-(t-t_1)/\tau_1) - 1) + c_2(\exp(-(t-t_1)/\tau_2) - 1)] \cdot \text{phase two } (t) \\ &+ [(A-a)(\exp(-(t-t_2) \cdot (A-a) \cdot s_{\text{consist}}) - 1) + A] \cdot \text{phase three } (t) \\ &\text{if condition} = \text{CONSISTENT} \end{aligned} \right\} \\ \left. \begin{aligned} &0 \cdot \text{phase one } (t) \\ &+ [c_3(\exp(-(t-t_1)/\tau_3) - 1) + c_4(\exp(-(t-t_1)/\tau_4) - 1)] \cdot \text{phase two } (t) \\ &+ [(B-a)(\exp(-(t-t_2) \cdot (B-a) \cdot s_{\text{revers}}) - 1) + B] \cdot \text{phase three } (t) \\ &\text{if condition} = \text{CUE REVERSAL} \end{aligned} \right\} \end{cases} \quad (1)$$

where

$$\begin{aligned} \text{phase one } (t) &= \begin{cases} 1 & \text{for } t < t_1 \\ 0 & \text{else} \end{cases} \\ \text{phase two } (t) &= \begin{cases} 1 & \text{for } t_1 \leq t < t_2 \\ 0 & \text{else} \end{cases}, \\ \text{phase three } (t) &= \begin{cases} 1 & \text{for } t_2 \leq t \\ 0 & \text{else} \end{cases} \end{aligned} \quad (2)$$

and

$$\begin{aligned} A &= c_1(\exp(-(t_2 - t_1)/\tau_1) - 1) + c_2(-(t_2 - t_1)/\tau_2 - 1) \\ B &= c_3(\exp(-(t_2 - t_1)/\tau_3) - 1) + c_4(-(t_2 - t_1)/\tau_4 - 1). \end{aligned} \quad (3)$$

$0 \cdot \text{phase one } (t)$ of course equals 0; we include the first term in the equations to show that the first leg is a zero line. We fixed t_2 to the value determined by the fit of FLIP, $t_2 = 96$ ms. The fit returned the following parameters: $c_1, c_2, c_3, c_4, a, t_1, \tau_1, \tau_2, \tau_3, \tau_4, s_{\text{consist}}, s_{\text{revers}}$. Parameters $c_1, c_2, c_3, c_4, \tau_1, \tau_2, \tau_3$, and τ_4 specify the sums of exponentials that describe the two signals in phase 2. a is the common asymptote for $t \rightarrow \infty$, and s_{consist} and s_{revers} are the initial slopes of the functions in phase 3. We used the inverse of the slopes to quantify persistence.

The equation for the FLIP condition was

$$\begin{aligned} y_{\text{flip}} &= 0 \cdot \text{phase one } (t) \\ &+ [c_5(\exp(-(t - t_5)/\tau_5) - 1) \\ &+ c_6(\exp(-(t - t_5)/\tau_6) - 1)] \cdot \text{phase two } (t) \\ &+ [\tau_7(s_{\text{flip}} - c_8/\tau_8)(\exp(-(t - t_6)/\tau_7) - 1) \\ &+ c_8(\exp(-(t - t_6)/\tau_8) - 1) + C] \cdot \text{phase three } (t), \end{aligned} \quad (4)$$

with

$$\begin{aligned} C &= c_5(\exp(-(t_6 - t_5)/\tau_5) - 1) \\ &+ c_6(\exp(-(t_6 - t_5)/\tau_6) - 1), \end{aligned} \quad (5)$$

and the equation for the ONSET condition was

$$\begin{aligned} y_{\text{onset}} &= 0 \cdot \text{phase one } (t) + 0 \cdot \text{phase two } (t) \\ &+ [\tau_9(s_{\text{onset}} - c_{10}/\tau_{10})(\exp(-(t - t_9)/\tau_9) - 1) \\ &+ c_{10}(\exp(-(t - t_9)/\tau_{10}) - 1)] \cdot \text{phase three } (t). \end{aligned} \quad (6)$$

The phase functions are defined as above, replacing t_1 and t_2 with t_5 and t_6 for FLIP and t_2 with t_9 for ONSET. The fit

for FLIP returned the parameters $c_5, c_6, c_8, t_5, t_6, \tau_5, \tau_6, \tau_7, \tau_8$, and s_{flip} and the fit for ONSET returned the parameters $c_{10}, t_9, \tau_9, \tau_{10}$, and s_{onset} . Parameters s_{flip} and s_{onset} are the initial slopes after flip and onset, respectively.

Results

The goal of this study was to see if the persistence of border ownership signals in V2 is an intrinsic property of the figure–ground mechanisms or if it reflects emerging object representations. Considering figure–ground organization as a process in a network of interconnected neurons, it is conceivable that, once the state of the network is set by the input signals, the network remains in that state until new signals arrive at the input that move it into a different state. Thus, if an edge is first assigned to object A, and then a new object B appears and new figure–ground cues indicate that the edge should now be assigned to B, the network will switch border ownership to B. In contrast, if the persistence of signals reflects object representations, we expect that an edge that is assigned to one object cannot easily be assigned to another object.

To distinguish between these hypotheses, we created displays in which the figure–ground cues reverse, while the objects remain the same. This was accomplished by displaying two figures, one occluding the other (Figure 1A, left), and then smoothly moving the bottom figure to a new position (Figure 1A, right). The final configuration is typically perceived as a square overlapping a rectangle. The black square that was initially in back now appears in front of the white figure. Thus, after the cessation of movement, the vertical edge in the center changes ownership from left to right. We refer to this condition as CUE REVERSAL (Movie S1).

For this test, the cue integration hypothesis predicts that the border ownership signal should change from “left” to “right” as soon as the figure–ground cues reverse. In contrast, the object representation hypothesis predicts persistence of signals, because the representation for the white L shape persists after the cessation of movement.

We tested this display sequence in neurons of area V2. The edge that underwent the change of border ownership was placed in the receptive field of the neuron under study (Figure 1A, red ellipse). For comparison, three other conditions were also tested: One was a similar movement display in which the other figure was in back and moved (Figure 1B). This produced a sequence in which the assignment of the border between the figures does not change (CONSISTENT, Movie S2). The other two conditions were: presentation of the final overlapping figure configuration without a history (ONSET, Figure 1C, Movie S3) and a figure–flip condition in which the light figure on the

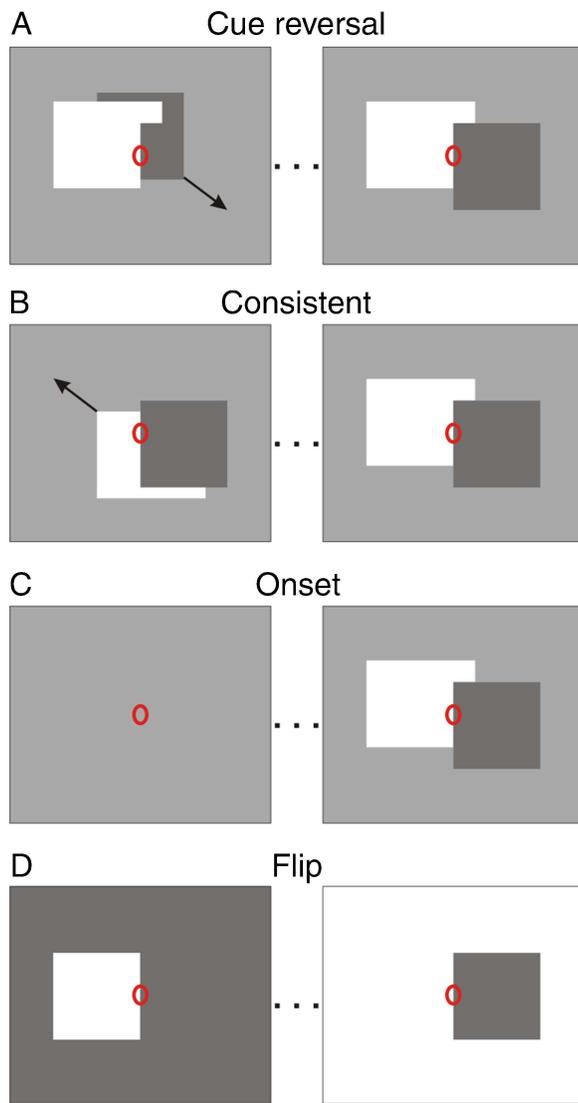


Figure 1. Displays for testing the dynamics of neural border ownership signals. (A) Reversal of figure–ground cues with object continuity. During the movement phase, the cues indicate that the border in the receptive field (red oval) is owned left, but in the final display they indicate ownership right. (B) Similar display sequence in which border ownership is “right” throughout. Note that the final configurations are identical in (A) and (B). (C) Presentation of a blank field followed by the final configuration. (D) Reversal of border ownership by simultaneously turning off a bright figure on the left and turning on a dark figure on the right. The actual test displays were rotated to match the preferred orientation of the neuron. Tests included displays with opposite border ownership and displays with reversed contrast.

left was deleted while a dark figure appeared on the right (FLIP, Figure 1D, Movie S4).

The comparison of the CONSISTENT and CUE REVERSAL conditions is shown in Figure 2. Figure 2A shows the signals of a sample neuron and Figure 2B shows the population averages. The CONSISTENT display

produced a sustained border ownership signal, as expected (blue traces). For the first 500 ms, foreground and background are defined by dynamic occlusion as well as geometric cues (shape and T-junctions). After the cessation of movement (time 0), only the geometric cues remain. In both phases, the cues indicate ownership right.

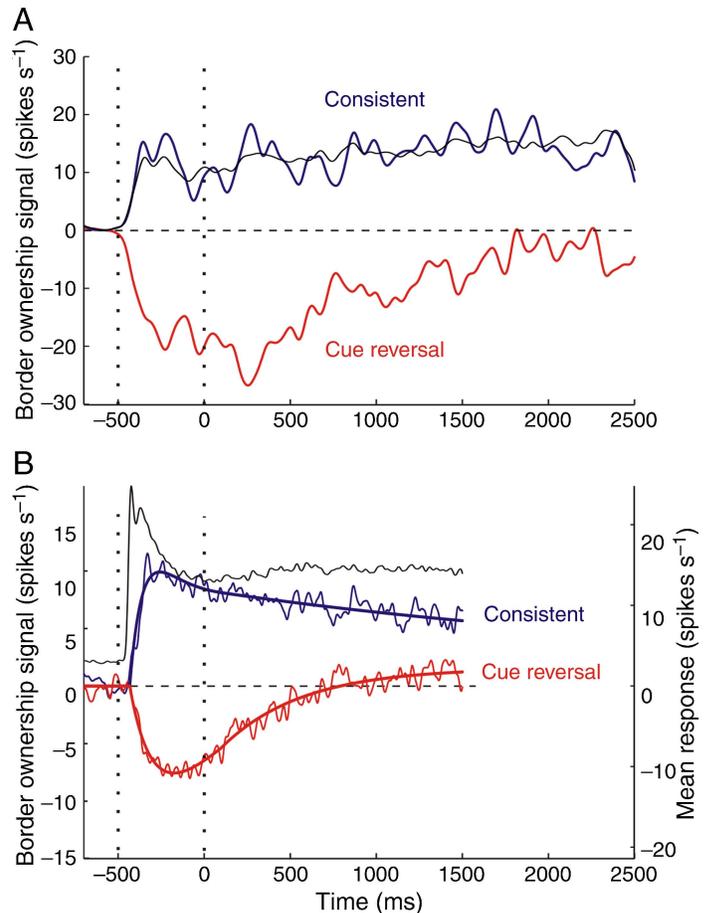


Figure 2. The neural border ownership signals for CUE REVERSAL and CONSISTENT display conditions. The border ownership signal is the difference between the responses to the display sequences ending in preferred and non-preferred border ownerships. (A) Sample neuron. (B) Means of 34 V2 neurons. The figures came on at time -500 and movement ended at time 0. During this phase, the figure–ground cues indicated “border ownership right” (corresponding to a positive signal) in CONSISTENT and “border ownership left” (corresponding to a negative signal) in CUE REVERSAL. After time 0, the displays were identical, showing “border ownership right,” but the signal in CUE REVERSAL remained negative for some time and then reversed slowly. Thin lines show smoothed histograms. Thin black line shows average of preferred and non-preferred side responses of both conditions. Thick lines show combinations of exponentials fitted to the data (see Materials and methods section). The traces for the sample neuron are longer than those for the population means, because this neuron was recorded using a longer fixation period (3.3 s) than the usual 2.3 s.

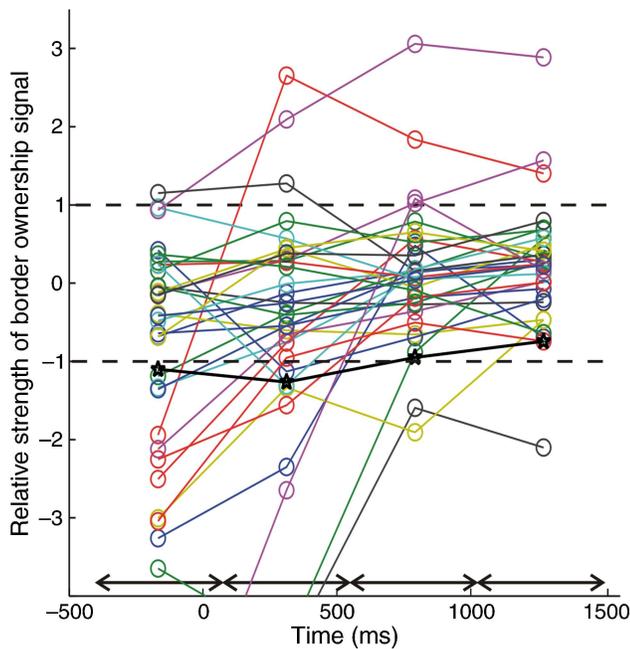


Figure 3. Persistence of the border ownership signals of the individual neurons in the CUE REVERSAL condition. For each neuron, the means of the signal during the four intervals shown by double arrows were calculated and normalized by its mean signal in the CONSISTENT condition. Circles connected by lines represent the individual neurons ($N = 34$); asterisks indicate the sample neuron of Figures 2 and 4. The first interval is the initial presentation phase in which overlay cues indicated “border ownership left” (as depicted in Figure 1), corresponding to a negative signal. After time 0, the CUE REVERSAL and CONSISTENT displays were identical, indicating “border ownership right”. Most neurons showed persistence of the negative signal, only slowly approaching the mean of the CONSISTENT condition (corresponding to the dashed line at 1).

The CUE REVERSAL display is similar, except that the cues in the motion phase indicate ownership left. Accordingly, the border ownership signal first goes negative (Figure 2, red traces). After the cessation of movement, the signal remains negative for about 700 ms despite the display now being identical to that of the CONSISTENT condition, indicating ownership right. The two signals slowly approach each other but do not reach a common level by the end of the fixation period. This shows that the initial border ownership assignment has a long-lasting effect despite the presence of new, contradictory figure–ground cues. The results were similar in the two animals (Supplementary Figure 1).

The variation between individual neurons is illustrated in Figure 3 (see Supplementary Figure 2 for the results separated by animal). The signals in CUE REVERSAL were averaged over four time bins, as indicated by double arrows on the time axis. Because the firing rates and hence the amplitudes of the border ownership signals

varied widely between neurons, the CUE REVERSAL signal of each neuron was normalized by the mean of its signal in the CONSISTENT condition. The figure shows that most neurons showed persistence of the negative signal, only slowly approaching the mean CONSISTENT level (represented at 1 in the graph because of the normalization). The median border ownership signals of CUE REVERSAL and CONSISTENT conditions were significantly different in each of the time bins ($p < 0.05$, Wilcoxon signed rank test).

The effect of stimulus history can be seen clearly by comparing the CUE REVERSAL with the ONSET condition (Figure 4, red and black traces): without history, the signal reaches its maximum value in less than 200 ms.

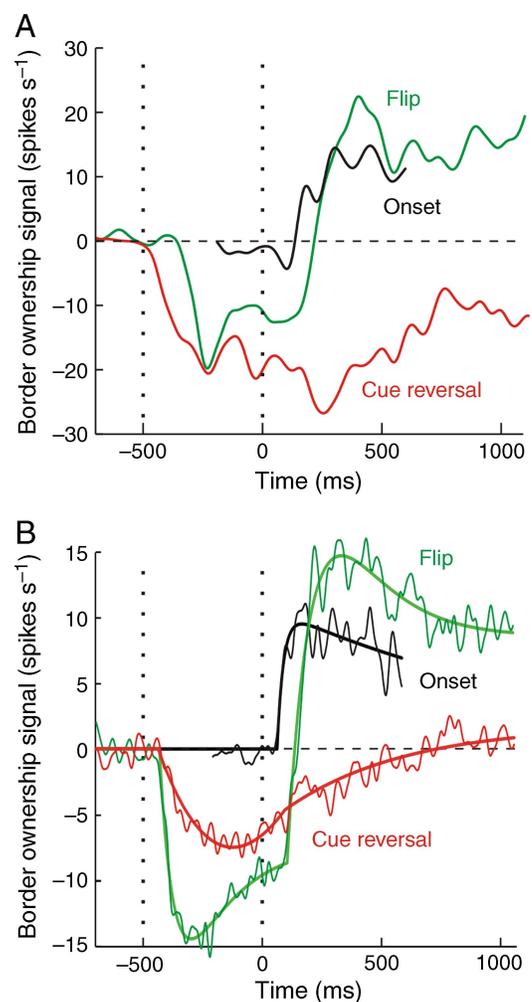


Figure 4. Comparison of the border ownership signals for CUE REVERSAL (red), immediate onset of the final configuration (ONSET, black), and the flipping figure condition (FLIP, green). (A) Sample neuron (same as in Figure 2). (B) Means of 27 V2 neurons. The FLIP signal is larger because isolated figures tend to produce stronger border ownership signals than overlapping figures. Note the fast rise of the signals in ONSET and FLIP compared to CUE REVERSAL.

Note that for the red and black curves, visual stimulation is identical after time 0.

In the FLIP condition, the border ownership cues reverse, as in CUE REVERSAL, but with the difference that the object of the first assignment is removed at the same time. The result was that the signal reversed quickly (Figure 4, green trace). Although the signal is first negative, as in CUE REVERSAL, the effect of replacing a figure on one side with a new figure on the opposite side is quite different. Note that the new figure in the FLIP condition is identical to the right figure in CUE REVERSAL at the end of movement. Thus, both conditions stimulated the same receptive fields, producing new edge signals at the same time in the same locations. The critical difference is that the figure to which the edge was initially assigned disappears in FLIP, whereas in CUE REVERSAL it continues to be visible.

Some other differences need to be considered too. (1) ONSET of course included the onset of an edge in the receptive field while this edge was turned on 500 ms earlier in the other conditions. The edge onset might explain some of the rapid change. Note, however, that FLIP produced a similar rapid change of the signal at time 0 without an edge transient. (2) The background changed in FLIP but not in CUE REVERSAL. However, the background change is unlikely to have a big influence because other experiments have shown that presentations of figures of the type of the Cornsweet illusion produce border ownership signals very similar to those produced by solid figures (Zhang & von der Heydt, 2010). In such a “Cornsweet figure,” the color/luminance varies only along a narrow seam at the contours. Thus, border ownership signals depend mainly on responses evoked by the contours. Moreover, we have shown that a change of background color does not interrupt the persistence of signals at an ambiguous edge (O'Herron & von der

Heydt, 2009; Figure 6). (3) The reason why FLIP produces a larger signal than CUE REVERSAL is that border ownership signals are larger for isolated figures than for borders between overlapping figures (Qiu, Sugihara, & von der Heydt, 2007).

The durations of signal persistence are summarized in Figure 5. There was a twentyfold difference in persistence between CUE REVERSAL and FLIP: When the owner of the edge disappeared (FLIP condition), the signal changed by 1 Hz in 3 ms, but when it continued to be visible (CUE REVERSAL), the same change took 65 ms. In the ONSET condition, the signal changed equally fast as in the FLIP condition. In the CONSISTENT condition, where occlusion cues continuously pointed in one direction, there was a slow negative signal change, indicating that the signal slowly adapted.

Discussion

We tested the hypothesis that border ownership signals reflect the persistence of emergent object representations. Such representations would enable tracking of object identity, facilitate selective processing, and provide the continuity needed for bridging cuts in the inflow of visual information caused by saccadic eye movements, blinks, or transient occlusions. We had recently found that border ownership signals persist when a figure display is switched to an ambiguous edge but can be “reset” by presentation of another figure (O'Herron & von der Heydt, 2009). We suggested that the signals persist because of the absence of figure–ground cues. We now find that, when objects in a display are rearranged so that figure–ground cues indicate a new assignment of an edge that was already assigned otherwise, the signals for the initial assignment persist despite the contradicting cues (Figure 2, CUE REVERSAL). This finding shows that the critical condition for the persistence is not the lack of figure–ground cues but a continued representation of the object to which the edge has been assigned.

The argument is based on a comparison of four conditions. Figure 2 shows the duration of the persistence. One and a half seconds after the displays were made identical, the signals of the CONSISTENT and CUE REVERSAL conditions still had not fully converged. These are the displays that portray object continuity. We know that the slow signal change in CUE REVERSAL is not due to inefficiency of figure–ground cues, because the signals rise quickly when the overlapping figures are presented directly (Figure 4, ONSET). In addition, the presence of a negative signal in itself does not hamper the transition to positive values: when an edge that was assigned to an object on one side is taken over by a new object on the other side, the signals change rapidly (Figure 4, FLIP).

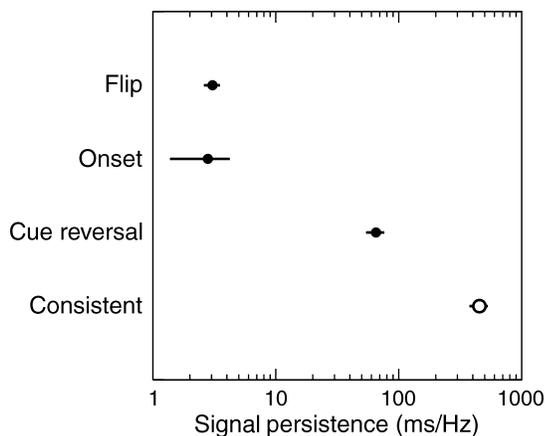


Figure 5. The persistence of border ownership signals in the four conditions illustrated in Figure 1. Persistence was defined as the inverse of the slope of the signal at the beginning of the final phase, as given by the function fits shown in Figures 2 and 4. The symbols represent the time (in milliseconds) it takes for the signal to change by 1 Hz. Open circle indicates negative value, that is, adaptation. Error bars represent SEM.

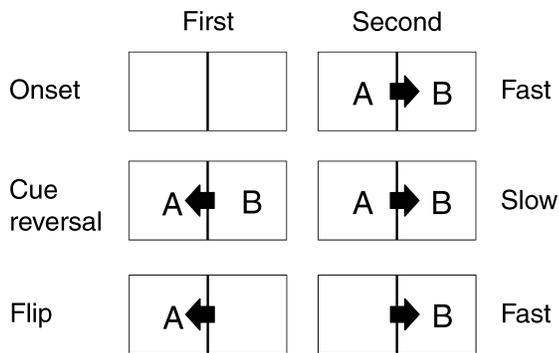


Figure 6. Schematic comparison of the signal dynamics in response to a change of border ownership cues under different conditions. Boxes mark the possible locations of objects in the two display phases, A and B indicate two different objects, and arrows indicate the direction of assignment according to the border ownership cues. The signal rises fast when objects appear without history (ONSET) but slowly when border assignment is reversed between existing objects (CUE REVERSAL). However, when the initial object of assignment disappears and a new object appears on the other side, the signal reverses fast (FLIP).

A schematic comparison of the conditions that produced the different signal transitions is shown in Figure 6. Boxes mark the possible object locations in the first and second phases of display, A and B stand for the two objects, and arrows indicate the direction of border assignment according to the figure–ground cues. The comparison between ONSET and CUE REVERSAL shows the effect of the stimulus history. The comparison between the CUE REVERSAL and FLIP conditions specifically shows the effect of object continuity. The main difference is that in CUE REVERSAL, the object to which the edge is initially assigned remained visible, whereas in FLIP it disappeared. When the object remained visible, the border ownership signal persisted 20 times longer than when the object disappeared.

We propose that the appearance of an object creates a representation in the visual cortex that has persistence. An important constraint is that border ownership signals emerge around 70 ms after stimulus onset (see Figures 2 and 3). This is before neurons in IT cortex become active (Bullier, 2001) and long before stimulus-triggered attention affects the activity in the visual cortex (Motter, 1994; Roelfsema et al., 1998; Super, Spekreijse, & Lamme, 2001). It means that the system does not use information from long-term object memory for the creation of these representations. A representation might consist in the activation of a node in a network that has connections to the neurons representing the object features: for example, a set of reciprocal connections between edge neurons in V2 and a small number of common “grouping cells,” where the grouping cells sum the edge signals from the figure and, by feedback, set the gain of the corresponding edge neurons. “Grouping node” might be a suitable term for such a circuit (for a

detailed description of such a model, see Craft, Schuetze, Niebur, & von der Heydt, 2007). This circuit produces border ownership selectivity in the edge neurons. It also enables selective enhancement of the responses of those neurons by top-down attention, which has been demonstrated experimentally (Qiu et al., 2007). To explain the present results, we assume (1) that grouping nodes are activated instantaneously by the stimulus but maintain their activity in the absence of sufficient input and (2) that there is mutual inhibition between grouping nodes. The persistence of activity in the grouping nodes explains the persistence of border ownership signals in CUE REVERSAL, and the mutual inhibition accounts for the rapid reversal of border ownership signals in the FLIP condition.

Is the continuity of representation a function of attention? The fact that we recorded persisting signals in monkeys that were trained to fixate and not saccade to the figures suggests that attention is not necessary. It is of course possible that the monkeys’ attention was drawn automatically to the figures, enabling persistence. However, experiments with sequential presentation of two figures ruled out this possibility (O’Herron & von der Heydt, 2009). The second figure was turned on 300 ms after the onset of the first figure, and after another 300 ms, both figures were reduced to ambiguous edges. If attention was drawn automatically to the figures, the onset of the second figure should have drawn attention away from the first figure. However, the persistence of the border ownership signals at the first figure was undiminished.

Persistence of visually evoked responses has generally been found only in experiments where the task required memorization of stimulus information. For example, the presentation of motion- or texture-defined figures produces enhancement of activity in V1 (Lamme, 1995), and this enhancement persists when a briefly presented figure is used as the target in a memory-guided saccade task (Super et al., 2001). The modulation persisted while the monkey attended to the location where the figure had been presented, but it decayed immediately when another target was presented to which the monkey had to saccade instead. Thus, in this case, the persistence depended on continued attention to the stimulated location. On the other hand, Lamme, Zipser, and Spekreijse (1998) found persistence of figure–ground modulation after brief presentation of motion-defined figures in the fixation paradigm, under conditions that are similar to those of the present experiments.

The question of whether visual cortical representations persist on their own or need attention to be maintained is of fundamental importance for understanding the function of the visual cortex. Our studies of border ownership signals indicate that temporary object representations (“grouping nodes”) are created automatically (Qiu et al., 2007) and persist without attention (O’Herron & von der Heydt, 2009). However, it is conceivable that, under certain task conditions, a change of attention might terminate the persistence. Clearly, further studies are needed to clarify

the influence of attention on the persistence and decay of figure–ground signals.

In conclusion, our results suggest that border ownership signals reflect the cortical representation of object continuity. Presumably, this representation plays a role in maintaining object identity across eye movements and object movements. Preliminary results support this prediction (O'Herron & von der Heydt, 2010).

Acknowledgments

This research was supported by NIH Grants EY02966 and EY016281 and ONR Grant N000141010278. We wish to thank Ofelia Garalde and Fangtu Qiu for technical assistance and Howard Egeth, Jonathan Flombaum, Anne Martin, and Ernst Niebur for comments on drafts of this paper.

Commercial relationships: none.

Corresponding author: Philip O'Herron.

Email: poherrol@jhmi.edu.

Address: Krieger Mind/Brain Institute, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA.

References

- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, *36*, 96–107.
- Craft, E., Schuetze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure–ground organization. *Journal of Neurophysiology*, *97*, 4310–4326.
- Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.
- Lamme, V. A. F. (1995). The neurophysiology of figure–ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605–1615.
- Lamme, V. A. F., Zipser, K., & Spekreijse, H. (1998). Figure–ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 3263–3268.
- Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of the primary visual cortex in higher level vision. *Vision Research*, *38*, 2429–2454.
- Motter, B. C. (1994). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience*, *14*, 2190–2199.
- O'Herron, P. J., & von der Heydt, R. (2009). Short-term memory for figure–ground organization in the visual cortex. *Neuron*, *61*, 801–809.
- O'Herron, P. J., & von der Heydt, R. (2010). Trans-saccadic memory for border ownership in neurons of the visual cortex. *Society for Neuroscience*.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking of multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 1–19.
- Qiu, F. T., Sugihara, T., & von der Heydt, R. (2007). Figure–ground mechanisms provide structure for selective attention. *Nature Neuroscience*, *10*, 1492–1499.
- Qiu, F. T., & von der Heydt, R. (2005). Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. *Neuron*, *47*, 155–166.
- Qiu, F. T., & von der Heydt, R. (2007). Neural representation of transparent overlay. *Nature Neuroscience*, *10*, 283–284.
- Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, *40*, 1469–1487.
- Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*, 376–381.
- Super, H., Spekreijse, H., & Lamme, V. A. F. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science*, *293*, 120–124.
- von der Heydt, R., Qiu, F. T., & He, Z. J. (2003). Neural mechanisms in border ownership assignment: Motion parallax and gestalt cues [Abstract]. *Journal of Vision*, *3*(9):666, 666a, <http://www.journalofvision.org/content/3/9/666>, doi:10.1167/3.9.666.
- Zhang, N. R., & von der Heydt, R. (2010). Analysis of the context integration mechanisms underlying figure–ground organization in the visual cortex. *Journal of Neuroscience*, *30*, 6482–6496.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*, 6594–6611.
- Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, *16*, 7376–7389.