

Spike Synchrony Reveals Emergence of Proto-Objects in Visual Cortex

 Anne B. Martin¹ and Rüdiger von der Heydt^{1,2}

¹Department of Neuroscience, the Johns Hopkins University, School of Medicine, Baltimore, Maryland 21205, and ²The Zanvyl Krieger Mind/Brain Institute, the Johns Hopkins University, Baltimore, Maryland 21218

Neurons at early stages of the visual cortex signal elemental features, such as pieces of contour, but how these signals are organized into perceptual objects is unclear. Theories have proposed that spiking synchrony between these neurons encodes how features are grouped (binding-by-synchrony), but recent studies did not find the predicted increase in synchrony with binding. Here we propose that features are grouped to “proto-objects” by intrinsic feedback circuits that enhance the responses of the participating feature neurons. This hypothesis predicts synchrony exclusively between feature neurons that receive feedback from the same grouping circuit. We recorded from neurons in macaque visual cortex and used border-ownership selectivity, an intrinsic property of the neurons, to infer whether or not two neurons are part of the same grouping circuit. We found that binding produced synchrony between same-circuit neurons, but not between other pairs of neurons, as predicted by the grouping hypothesis. In a selective attention task, synchrony emerged with ignored as well as attended objects, and higher synchrony was associated with faster behavioral responses, as would be expected from early grouping mechanisms that provide the structure for object-based processing. Thus, synchrony could be produced by automatic activation of intrinsic grouping circuits. However, the binding-related elevation of synchrony was weak compared with its random fluctuations, arguing against synchrony as a code for binding. In contrast, feedback grouping circuits encode binding by modulating the response strength of related feature neurons. Thus, our results suggest a novel coding mechanism that might underlie the proto-objects of perception.

Key words: area V2; attention; binding; figure–ground organization; objects; spike synchrony

Introduction

An unresolved question in neuroscience is how the brain represents objects. The early stages of the visual cortex, areas V1 and V2, each containing hundreds of millions of visually selective neurons, hold the most detailed information. But these neurons have small receptive fields representing elemental features. How these features are grouped into objects is not clear. Neurons in the inferotemporal cortex have larger receptive fields and can signal combinations of local features as needed for object recognition, but this seems to rely on an earlier stage of large-scale organization. In cluttered scenes where objects occlude one another, as in ordinary visual environments, object recognition and search depend on the identification of foreground and background and the correct assignment of contours to the foreground (border-ownership assignment; Nakayama et al., 1989, 1995; Rensink and Enns, 1998; Baylis and Driver, 2001). It has therefore been sug-

gested that a stage of automatic organization first creates a preliminary representation of objects (“figures” as distinct from ground, “object files,” or “proto-objects”; Koffka, 1935; Kanizsa, 1979; Kahneman et al., 1992; Rensink, 2000; Pylyshyn, 2001) or object surfaces (Nakayama et al., 1995).

If elemental features are represented by the firing rate code, then representation of objects, which are conjunctions of features, seems to require an independent code, and this could be spiking synchrony between neurons (binding-by-synchrony; von der Malsburg, 1986; Gray et al., 1989). However, recent studies did not find the predicted synchrony (Lamme and Spekreijse, 1998; Roelfsema et al., 2004; Palanca and DeAngelis, 2005; Chen et al., 2014). Moreover, studies of figure–ground organization in visual cortex have shown that the firing rates of feature responses often depend on the image context (Lamme, 1995; Bakin et al., 2000; Duncan et al., 2000; Zhou et al., 2000; Li et al., 2006; Huang et al., 2008; Zhang and von der Heydt, 2010), indicating that rate modulation may also code for the assignment of features to objects. For example, some neurons give differential responses to an edge depending on whether it is the border of a figure on one side or the other (“border-ownership selectivity”; Zhou et al., 2000). Because this selectivity is a fixed property of the neurons, it has been proposed that feature neurons are connected to “grouping cells” at a higher level that facilitate their responses by feedback (Fig. 1A; Craft et al., 2007; Qiu et al., 2007).

Received Aug. 27, 2014; revised March 24, 2015; accepted March 26, 2015.

Author contributions: A.B.M. and R.v.d.H. designed research; A.B.M. performed research; A.B.M. and R.v.d.H. analyzed data; A.B.M. and R.v.d.H. wrote the paper.

This work was supported by National Institutes of Health Grants R01-EY002966 and R01EY016281 and Office of Naval Research Grant N000141010278. We thank Ofelia Garalde and Fangtu T. Qiu for technical assistance, and Ernst Niebur and Robert Desimone for critical comments on previous versions of this manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Anne B. Martin, Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544. E-mail: abmartin@princeton.edu.

DOI:10.1523/JNEUROSCI.3590-14.2015

Copyright © 2015 the authors 0270-6474/15/356860-11\$15.00/0

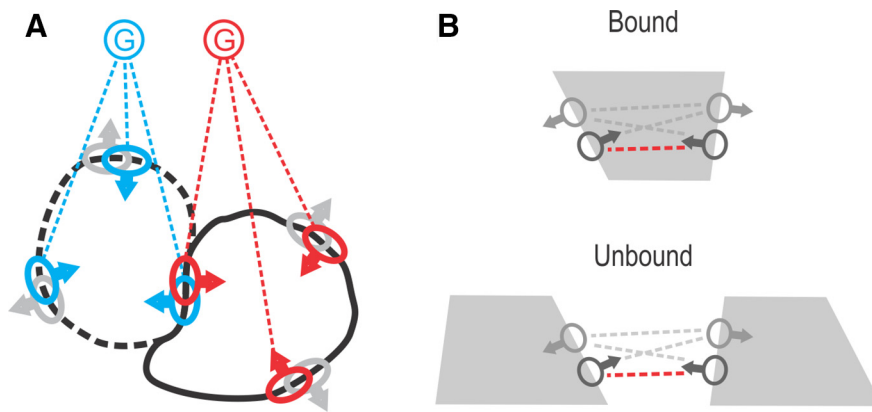


Figure 1. The hypothesis of feedback grouping circuits. **A**, Local-feature-selective neurons, in addition to projecting to downstream processing areas (not shown), send collaterals to grouping cells (G) that enhance their activity through feedback connections (red and blue dashed lines). This feedback makes the neurons selective for the side of border ownership, as indicated by arrows on the receptive field symbols (ellipses). **B**, The hypothesis predicts (1) that pairs of neurons whose border-ownership preferences are consistent with a common object (symbols connected by red dashed lines) will have increased synchrony when stimulated by a common object (bound) compared with two separate objects (unbound), while (2) pairs of neurons with inconsistent preferences (symbols connected by gray dashed lines) will show no synchrony, regardless of the binding condition.

A critical prediction of this hypothesis is that feature neurons connected to the same grouping cell will fire more synchronous spikes than other neurons when that grouping cell is activated, because they receive the same spike trains in the feedback. The main goal of this study was to test this prediction. Moreover, since proto-objects are the units of perception, the strength of grouping should facilitate selective attention. Thus, we also examined whether the degree of synchrony, which reflects the strength of grouping feedback, correlates positively with performance in selective shape discrimination. Our results provide evidence for grouping feedback as the mechanism of proto-object formation.

Materials and Methods

All procedures conformed to the guidelines of the National Institutes of Health as approved by the Johns Hopkins University Animal Care and Use Committee.

Recording. Using standard surgical and electrophysiological techniques (Zhou et al., 2000), we simultaneously recorded well-isolated extracellular action potentials from two independently controlled microelectrodes. We used electrodes with fine tips, which easily isolated single units, but also picked up some background activity (platinum–iridium alloy; 0.1 mm diameter; etched taper, ~10%; glass coated; impedance, 2–9 M Ω at 1 kHz). The two electrodes were inserted parallel to each other, separated by a distance of 3–14 mm. We recorded cells from V1 and V2 (in the lip of the postlunate gyrus or in the lunate sulcus after passing through V1 and the white matter) from four hemispheres in three male rhesus macaques (*Macaca mulatta*). The signal was split to separately extract the local field potential (LFP) and spike components. The spike signal was bandpass filtered at 250 Hz–6 kHz (24 dB/octave) and digitized at 50 kHz. The LFP signal was bandpass filtered at 0.7–250 Hz, and digitized at 1 kHz. Spikes were isolated using the Alpha Omega spike detection system (ASD 2.80). The spike time data here show the times of action potentials only for well isolated spikes, excluding noise and multiunit activity. Occasionally, two clear waveforms were apparent on the same channel. In those cases, their spike times were separated. The powerline artifact was not removed from the LFP since it did not affect the results (analyses of LFP signals with the line-noise artifact digitally removed were not different from those with the artifact present). The direction of gaze was monitored for one eye using an infrared video-based system (Iscan ETL-200) at 60 Hz with angular resolution of 0.08° visual angle horizontally and 0.16° vertically. Spike times, spike shapes, LFPs, stimulus events, behavioral events, and gaze directions were recorded using VLab, a custom system (Fangtu T. Qiu, personal communication) for off-line analysis.

Receptive field positions, preferred orientations, and border-ownership selectivity were mapped using standard techniques previously reported (Qiu and von der Heydt, 2005).

Visual stimuli and experimental design. Stimuli were generated using Open Inventor and were presented on either a 21 inch Eizo FlexScan T965 or ViewSonic G220fb color monitor with 1600 \times 1200-pixel resolution at 100 Hz refresh rate. The field of view subtended a 20° by 15° visual angle viewed at a distance of 1 m. A neutral gray background of 26 cd/m² luminance was used.

In the main experiment, three figures were presented in two configurations (Fig. 2A). The receptive fields of the mapped neurons were used to define the location and shape of the figures. In the “bound” configuration, one quadrilateral figure was presented with opposing edges in the two receptive fields at each neuron’s preferred orientation, with the other edges parallel, thereby forming a trapezoid. The other two figures were of the same shape, but horizontally displaced, generally in the other visual hemifield, and rotated such that the gap between them matched the shape of the bound figure. In the “unbound” configuration, the two edges in the neurons’ receptive field were identical to those in the bound configuration, except that in this case they were part of separate figures, and the third figure was the same shape horizontally offset. Thus all three figures in both configurations had comparable eccentricities. The edges in the receptive fields were $\geq 50\%$ longer than the diameter of the corresponding classical receptive field, extending on both sides, so that any context information about the figure was outside the classical receptive field. Thus, each neuron was stimulated by a straight contour at its preferred orientation. If the configuration of receptive field positions and orientations made it impossible to create a trapezoid shape between them, then such pairs were not tested.

Each display consisted of two colors, one for the figures and one for the background. One color was chosen to match the neurons’ color preference if there was any, or otherwise white. The other color was a medium gray. The assignment of the colors to figures and background was permuted between trials as shown in the four example displays of Figure 2A. The color reversal in combination with the two configurations enabled us to separate the effects of border ownership and binding from effects of local edge contrast (Zhou et al., 2000).

In each fixation trial (see below) one of the two configurations and one of the two contrast polarities was presented, and one of the three figures was assigned the target of attention (the manipulation of attention is described in Behavioral task below). Each of the three figures could be the target. Thus, the binding and contrast polarity variables each had two levels, while the attention variable had three levels nested within binding (because binding changed the figure arrangement). For the analysis of synchrony and coherence we excluded trials of the unbound conditions in which a figure at one of the receptive fields was the target (Fig. 2A, bottom), because the hypotheses we tested make no predictions for this case, but all conditions were used to calculate the border-ownership preferences of the neurons.

Trials were organized in blocks of 8 or 16. Within a block, the configuration of shapes and target assignment did not change while the local contrast polarity and response type (lever press or pull; see Behavioral task) were pseudorandomly varied and balanced over the block. Shape configuration and target figure were pseudorandomly reassigned between blocks and balanced over the set of blocks.

Behavioral task. The animals were trained to perform two kinds of tasks as signaled to them by the shape of the fixation spot. The first was a fixation task that was used for mapping receptive fields and preliminary analysis. It required the animals to maintain the gaze direction signal within a 1° radius window centered on the fixation spot. For the main test

of our experiment, the animals performed a shape discrimination task in which they were required to indicate by manipulating a joystick whether the target figure changed shape or moved (Fig. 2B), while ignoring any changes of the distracter figures. Throughout, they were required to maintain the gaze direction signal within a 1° radius window centered on the fixation spot. In instruction trials before each block of trials, one of the figures was designated as the target for the behavioral task by blinking, and this figure determined the reward contingency throughout the block. Instruction trials were excluded from all analyses.

The shape change or movement of the target and distracter figures was produced by moving the two opposing edges of each trapezoid. In each trial, four moving edges (two of the target and two of the distracter) moved independently and pseudorandomly, creating shape change and movement with equal frequency (for monkey GW, only the target underwent a change because he was not able to learn the task at above-chance levels when distracters also changed). The distance of edge movement was limited by the size of the figure, since in the case of a contraction the edges come together; expansion and movement distances were set equal to the contraction distance; the edges typically moved beyond the extent of the recorded receptive fields before the monkey responded; however, because we analyzed synchrony only during the static period, i.e., before the edges begin to move, the extent of the movement is not relevant to the analysis.

In a successful trial (Fig. 2C), a central fixation spot appeared after an intertrial interval with a blank screen (1000 ms). The animals found and fixated on the spot for 300 ms, after which time the three figures were displayed. After 1000 ms, the target figure and one of the distracters began to change shape or move, and the animal had to respond according to the target figure (for monkey GW, the onset of the critical change was delayed randomly by 500–1500 ms; from these data, only trials with movement after 1000 ms were included for analysis). If the monkey responded correctly, he received a drop of juice reward, and the screen cleared for the intertrial interval of 1 s. In incorrect trials, or in the case of fixation breaks, the animal received no juice reward and the intertrial interval was increased by 1 s. A single cue trial was inserted if the monkey performed >3 consecutive incorrect trials. Cue trials were not included in the analysis.

Spike train and LFP analysis. The analysis of spike and LFP recordings focused on the interval 400–1050 ms after figure onset where the mean firing rate was stationary, excluding the initial transient of the responses. Neuron pairs where either neuron was not sufficiently driven by the stimuli (≥ 5 spikes/s mean firing rate in at least 1 condition) or pairs with <8 completed trials per condition were excluded.

Border-ownership selectivity was determined by a three-way ANOVA on the square-root transformed spike counts with binding, attend, and contrast polarity as independent variables. The square-root transformation $x \rightarrow (x + 3/8)^{1/2}$ (Bartlett, 1936) was used to stabilize the variance (the variances of the spike counts of repeated responses typically vary in proportion to their mean; Vogels et al., 1989).

Covariograms and spike rate correlations. Covariograms and trial-by-trial spike rate correlations (“noise correlations”) between simultaneously recorded spike trains were computed separately for each of the conditions given by the variables local contrast, binding, and attention.

Noise correlations were calculated by finding, for each condition and pair, the Pearson’s correlation coefficient of the square-root transformed spike counts in the stationary period (400–1050 ms after figure onset).

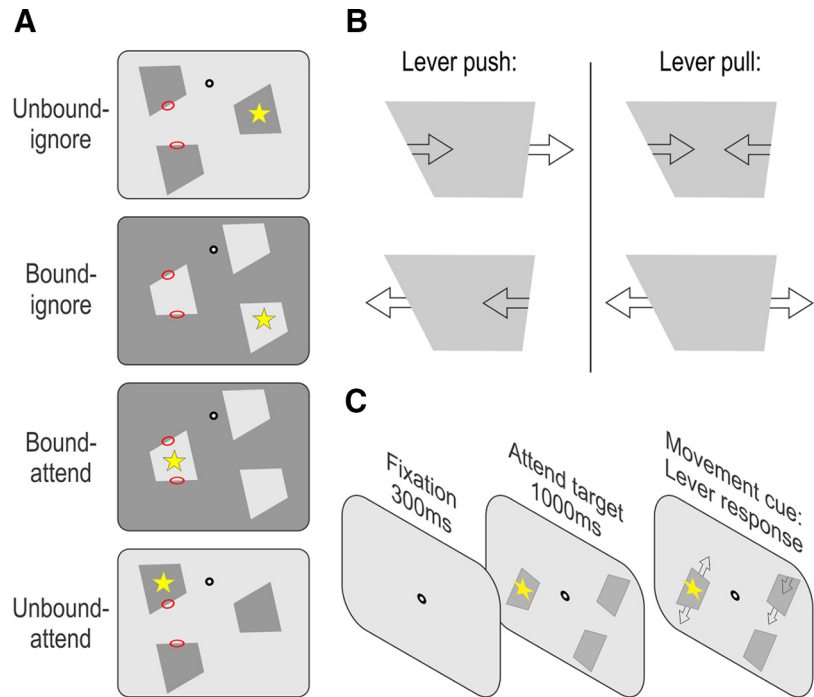


Figure 2. Experimental design. **A**, Schematic illustration of the four basic conditions of the experiment. Three trapezoidal figures were presented simultaneously so that two edges activated the receptive fields of the recorded neurons at their preferred orientations (red ellipses, not part of the display). The two edges were either part of two separate figures (unbound) or of the same figure (bound). The contrast was reversed between bound and unbound conditions to make the edges in the receptive fields identical. Both conditions were also tested with the opposite contrast (not shown). The monkey attended one of the figures as indicated by a yellow star (not part of the display) according to instruction given before a block of trials. **B**, The monkey is rewarded if he correctly indicates by lever movement whether the edges move in phase, producing object movement (left), or in antiphase, producing contraction or expansion (right). **C**, Task. Monkey initiates a trial by fixating a central point. Then the three figures appear, one of which has been cued as the target. After 1000 ms, opposite edges in the target and one of the distracter figures begin to move and the monkey responds using a lever.

Covariograms were calculated either for the entire stationary period (400–1050 ms after figure onset) or in sliding windows 250 ms wide in 10 ms steps. The spike time records were converted to time series with 1 ms resolution, assigning a value of 1 if a spike was recorded and zero otherwise, in the following called spike trains. The covariogram between spike trains S_j, S_k from neurons j and k was computed as follows:

$$C_{j,k}(\tau) = \frac{\langle (S_j^i - e_j^i) \odot (S_k^i - e_k^i) \rangle_i(\tau) - \langle S_j^i - e_j^i \rangle_i(\tau) \langle S_k^i - e_k^i \rangle_i(\tau)}{\Theta(\tau)}$$

where $\langle X \rangle_i$ indicates the mean across trials,

$$S_j^i \odot S_k^i(\tau) = \sum_{\mu=t_0-\tau}^{t_{end}+\tau} S_j^i(\mu + \tau) S_k^i(\mu)$$

is the cross-correlation function between the i^{th} trial of spike trains $S_j^i, S_k^i(w)$ is the half-width of the window of cross-correlation, t_0 and t_{end} define the analysis period),

$$e_j^i = \frac{1}{t_{end} - t_0} \sum_{n=t_0}^{t_{end}} S_j^i(n)$$

is the mean spike count per bin in trial i of train j , and $\Theta(\tau) = t_{end} - t_0 - |\tau|$ is a triangular function used to correct for the varying amount of overlap between the two spike trains at each time lag τ .

Because the spike trains are time-density functions (e.g., counts/ms) the covariogram has the dimension of coincidences/s². To represent the population, covariograms were averaged across neuron pairs within the pair groups and conditions after symmetrizing each covariogram according to the following equation: $C_{sym}(\tau) = \frac{1}{2}[C(\tau) + C(-\tau)]$.

Synchrony. Synchrony was calculated by integrating the covariograms over a certain interval around zero as follows:

$$Y_{j,k} = \sum_{\tau=-\lambda}^{\lambda} C_{j,k}(\tau) \cdot \text{binsize}$$

This gives us the frequency of coincident spikes in cells j and k in excess of the level expected by chance. We will include λ in statements about the degree of synchrony. For example, by “40 ms synchrony,” we mean the frequency of coincidences obtained by integrating the covariogram over the range ± 40 ms. Because the covariograms correct for effects of firing rate variation, the synchrony thus calculated should correctly be termed “excess synchrony,” synchrony that exceeds the rate of coincidences expected by chance. However, for brevity we will generally call it just synchrony.

The overall mean level of synchrony expected by chance was calculated from the average across conditions of the cross-correlation function of the peristimulus time histograms.

To test for precise synchrony we used an interval jitter method (Amarasingham et al., 2012). This method tests the specific null hypothesis that there was no temporal structure on a scale finer than a chosen bin width δ . For a given δ , surrogate spike trains are created from each original spike train by randomizing spike times with a uniform distribution in bins of length δ . This procedure maintains the coarse firing rate profile of each response at the resolution defined by δ . Repeating this jittering 1000 times produced ensembles of surrogate spike trains from which a distribution of surrogate cross-correlograms was calculated. The mean of this distribution was subtracted from the cross-correlogram of the original spike trains, as well as from each surrogate cross-correlogram, providing a jitter-derived covariogram and confidence limits of the null hypothesis.

Coherence. Coherence between pairs of spikes, or between spikes and LFPs, was measured in the period from 400 to 1050 ms after stimulus onset using the multitaper method (Mitra and Pesaran, 1999; Jarvis and Mitra, 2001), with functions modified from the Chronux data analysis toolbox for Matlab (www.chronux.org). The cross-spectrum S_{xy} is the mean over tapers (k) and trials (N) of the Fourier transform of one signal, x , times the complex conjugate of the Fourier transform of the other signal, y : $S_{xy}(f) = \langle J_x(f)J_y^*(f) \rangle_{N,k}$ where

$$J_x(f) = \sum_{t=1}^T w_k(t) x_t e^{-2\pi i f t}$$

Coherence is the magnitude of the cross-spectrum of two signals divided by the geometric mean of the two spectra:

$$C_{xy}(f) = \left| \frac{S_{xy}(f)}{\sqrt{S_x(f)S_y(f)}} \right|$$

and has values between 0 and 1, where signals having a low covariation of amplitude or variable phase relation have coherence close to zero, and signals with high amplitude covariation and constant phase relation will have a coherence close to 1. For spike–spike coherence, we used 20 Slepian tapers to get an effective spectral smoothing of ± 16 Hz at low frequencies ($f < 30$ Hz) and 38 tapers for ± 30 Hz smoothing in gamma and high gamma ($f \geq 30$ Hz). For spike-field coherence, we used nine tapers for ± 8 Hz smoothing at low frequencies and 20 tapers for ± 16 Hz smoothing in gamma and high gamma. Spike and LFP signals were mean-free, and LFP signals were detrended before coherence analysis. Coherence has a bias where it approaches 1 as the number of trials goes to 1, but this bias becomes negligible at trial counts > 10 . We considered a correction for this bias using the transformation in Bokil et al. (2007), but the correction did not affect our results since our data had ≥ 8 trials per condition and typically had many more; the correction is not applied in the results reported here.

Covariograms and coherence spectra were smoothed (using a Gaussian kernel of $\sigma = 4$ ms for covariograms, $\sigma = 4$ Hz for coherence) for clarity of the display.

Behavioral correlations. To determine whether synchrony was related to reaction time, we performed two analyses. First, we computed, over all pairs

and trials in each condition and pair group, the regression of reaction time on synchrony. Because reaction time has a distribution that deviates from normality, we applied the transformation $RT_i = \sqrt{RT - 300}$ ms, which produced an approximately normal distribution. Next, for each pair and condition we found the mean synchrony in the quartiles of trials with the fastest and slowest reaction time. We compared the population means of consistent and inconsistent pairs (as defined in Results) of the bound and unbound conditions in these quartiles.

Significance testing. We tested whether the amount of synchrony or coherence in two conditions was significantly different by randomly assigning trials to conditions, maintaining the number of trials per condition, and calculating the difference between the means, repeating the whole procedure 1000 times. The quantile of the actual difference in the resulting “null distribution” of randomized differences is the p value of the test. We used a similar randomization test to determine the significance of differences between the groups of pairs (consistent and inconsistent), where the null distribution was formed by randomly assigning pairs to groups.

Analysis of eye movements. The possible influence of eye movements on the results was determined by analyzing (1) the mean gaze position during the static stimulus periods and (2) the frequency of microsaccades during these periods. Gaze position was examined by ANOVA with the experimental variables binding, attend, and contrast polarity as factors. The frequency of microsaccades was examined by two separate ANOVAs, one for the ignore condition (attend = 0) with binding and contrast polarity as factors, and a second one for the bound condition (binding = 1) with attend and contrast polarity as factors.

Results

We evaluated our results with regard to three hypotheses, the binding-by-synchrony hypothesis (von der Malsburg, 1986; Gray et al., 1989), the attention coding hypothesis (Niebur and Koch, 1994), and the grouping hypothesis outlined here. For the binding-by-synchrony hypothesis, the critical question is whether neurons exhibit a higher degree of synchrony when they are stimulated by the same object (the binding condition) than when they are stimulated by different objects. Specifically, the added synchrony should be sufficiently strong relative to the variability of synchrony across recording sites to provide a robust population signal for feature grouping. The attention coding hypothesis faces a similar question: does attention increase synchrony sufficiently so that downstream centers can selectively process the features of the attended object.

The critical question for the grouping hypothesis is whether binding produces a greater increase of synchrony between feature neurons in the same grouping circuit compared with feature neurons in different grouping circuits. Figure 1 shows that whether two neurons are part of the same grouping circuit is related to their border-ownership preferences: pairs of neurons with border-ownership preferences pointing toward each other (consistent pairs; Fig. 1B, red dashed lines) likely receive common grouping inputs, and neurons with inconsistent preferences (Fig. 1B, gray dashed lines) likely receive disparate grouping inputs. Note that in this theory, synchrony demonstrates a specific neural connectivity, and it is this connectivity that underlies the enhancement of feature responses with binding and attention. The amount of synchrony is not critical because selective processing is based on the enhancement of firing rates (for which there is evidence; Reynolds et al., 2000; Roelfsema et al., 2004; Qiu et al., 2007; Wannig et al., 2011), although it may be further facilitated by increased synchrony.

We recorded pairs of neurons from separate microelectrodes in macaque visual cortex (areas V1, V2) of three monkeys while varying the stimulus configuration so that the neurons were stimulated either by the same object (bound) or by separate objects

(unbound; Fig. 1B). The objects were shaped with contours in the neurons' receptive fields such that each was stimulated at its optimal orientation.

To see the influence of attention, we trained the monkeys to covertly attend one of three simultaneously presented figures in anticipation of a response cue (Fig. 2A). To obtain a reward, they had to discriminate whether the target figure changed shape or moved. Because we wanted the system to process the distributed feature signals of an object as a whole, we designed the task so that the animal had to use the edge signals from opposite sides of the target figure. Each trial started with a static presentation of the three figures (1000 ms), after which two edges of the target moved either in phase, producing object movement, or in antiphase, producing expansion or contraction (Fig. 2B,C). Similar but independent movements were applied to the edges in one of the distractor figures in two of the three monkeys (see Materials and Methods). To obtain a reward, the animal had to correctly indicate with a lever whether the target changed shape or moved, while maintaining central fixation throughout (1° radius fixation window). A block design was used and configuration and designation of the target remained the same throughout the blocks (8 or 16 trials). We analyzed spike synchrony in three conditions (unbound-ignore, bound-ignore, bound-attend) during the static period when the animal covertly attended the target figure (the responses to the edge movements were not included). The bottom condition in Figure 2A (unbound-attend) was not used in the analysis of synchrony because the hypotheses make no clear prediction for this situation, but all conditions were used for determining border-ownership selectivity and attentional response enhancement.

Of 190 pairs recorded, 104 had sufficient trials for analysis (minimum eight trials, maximum 394, mean 80.6) in every condition (monkey GW, 15; monkey DW, 41; monkey BE, 48). We classified 38 pairs as consistent (GW, 8; DW, 15; BE, 15) and 66 as inconsistent based on the difference in firing rates in response to a figure on one side compared with the other (preferred side produced 32% stronger responses on average; CI, 27–38%; $N = 208$). Attention produced 5.5% enhancement of responses in consistent pairs in the bound condition ($N = 76$; CI, 1.5–9.4%; $p < 0.008$, paired t test). We found no differences in synchrony between area classifications of the pairs (V1–V1, $N = 25$; V1–V2, $N = 29$; V2–V2, $N = 42$; not classified, $N = 8$; $p = 0.14$, ANOVA with factors pair consistency type, condition, area classification) or between the three monkeys ($p = 0.08$, ANOVA with factors pair consistency type, condition, monkey). We therefore pooled the data from the three monkeys and the pairs from different areas.

The reaction times were comparable between the monkeys (monkey GW: mean reaction time, 446 ± 72 ms; accuracy, 98.5%; monkey DW: mean reaction time, 504 ± 87 ms; accuracy, 82.3%; monkey BE: mean reaction time, 422 ± 77 ms; accuracy, 70.8%). In each monkey, reaction times did not vary across conditions and pair consistency types (condition $p = 0.11$, pair consistency type $p = 0.9$, ANOVA with factors pair consistency type, condition, monkey). Thus all conditions were equally difficult for each monkey and their reaction times did not change between periods when consistent and inconsistent pairs were recorded. Because one monkey (GW) was not able to perform the task above chance level with distracter movement, he was tested with only target movement, which resulted in higher overall accuracy than the other monkeys, but his neurophysiological results were nevertheless consistent in all respects with those of the others.

Because the figures were shaped based on the recorded neurons' receptive field location and preferred orientation, each pair led to the creation of a different configuration of figures. It was interesting to note that the animals were able to generalize the task to any trapezoid shape we presented. Receptive field positions were in the lower hemifield at eccentricities of 0.3 – 3.1° (mean 1.5°) for the more central receptive field of a pair, and 1.1 – 4.3° (mean 2.6°) for the more peripheral receptive field. The figures varied in size according to the distance between the receptive fields (0.6 – 3.8° ; mean, 1.8°) and distance from fixation (1.0 – 3.4° ; mean, 2.1°), but in all pairs, the receptive fields were exclusively nonoverlapping, and there was no difference in the size of figures defined by consistent (0.6 – 3.8° ; mean, 1.8°) or inconsistent pairs (0.6 – 3.3° ; mean, 1.8°).

Effects of binding and attention on synchrony

We measured spike synchrony by calculating cross-correlations between simultaneously recorded spike trains and correcting for coincidences expected by chance (see Materials and Methods). We subtracted individual shuffle predictors from the cross-correlograms for each of the binding and attention conditions. The shuffle predictor, which is the cross-correlation function of the two peristimulus time histograms, is designed to correct for the stimulus-locked covariance of firing rates. However, in the presence of common fluctuations of excitability in the two neurons, the shuffle predictor is inaccurate, which can lead to peaks in the corrected cross-correlogram even when the spike trains are completely independent (Brody, 1999). To eliminate this possibility we subtracted from the spike train of each single response the mean firing rate of that response during the analysis period before calculating the cross-correlation function and shuffle predictor, as suggested by Roelfsema et al. (2004). As a further precaution, we excluded from the analysis the initial part of the responses where firing rates vary rapidly, using only the stationary period from 400 to 1050 ms after stimulus onset (as a result, the shuffle predictors were essentially flat). We refer to these corrected cross-correlation functions as covariograms (see Materials and Methods). Excess synchrony (i.e., synchrony beyond what could be expected by chance) was determined by integrating the covariogram over a chosen interval about zero.

We found that consistent pairs exhibited synchrony that varied with object configuration and attention (Fig. 3). Covariograms of example pairs from each monkey (Fig. 3B) show the excess coincidences as a function of time lag between spikes and time across the trial. The effect of binding can be seen by comparing the bound-ignore and unbound-ignore conditions (middle and top rows). The covariograms show clear positive peaks centered on zero lag that last throughout the trial in the bound configuration, compared with generally low coincidence counts in the unbound configuration, indicating increased synchrony with binding. When the binding figure was attended, synchrony was lower compared with when the figure was ignored (compare bottom and middle rows).

The same effects can be seen in the population averages (Figs. 3C, 4). To calculate the average, the covariograms of the individual neuron pairs were symmetrized assuming that the order of neurons in a pair is arbitrary. Indeed, we found no systematic deviations of the peaks from zero even among the V1–V2 pairs. Correlations between consistent pairs peak at zero lag, and binding increases the amplitude of the peak (Fig. 4A, left, black solid vs dashed lines), while attention decreases it (solid yellow vs black lines). For the inconsistent pairs, correlations are low and relatively broad in all conditions (Fig. 4A, right).

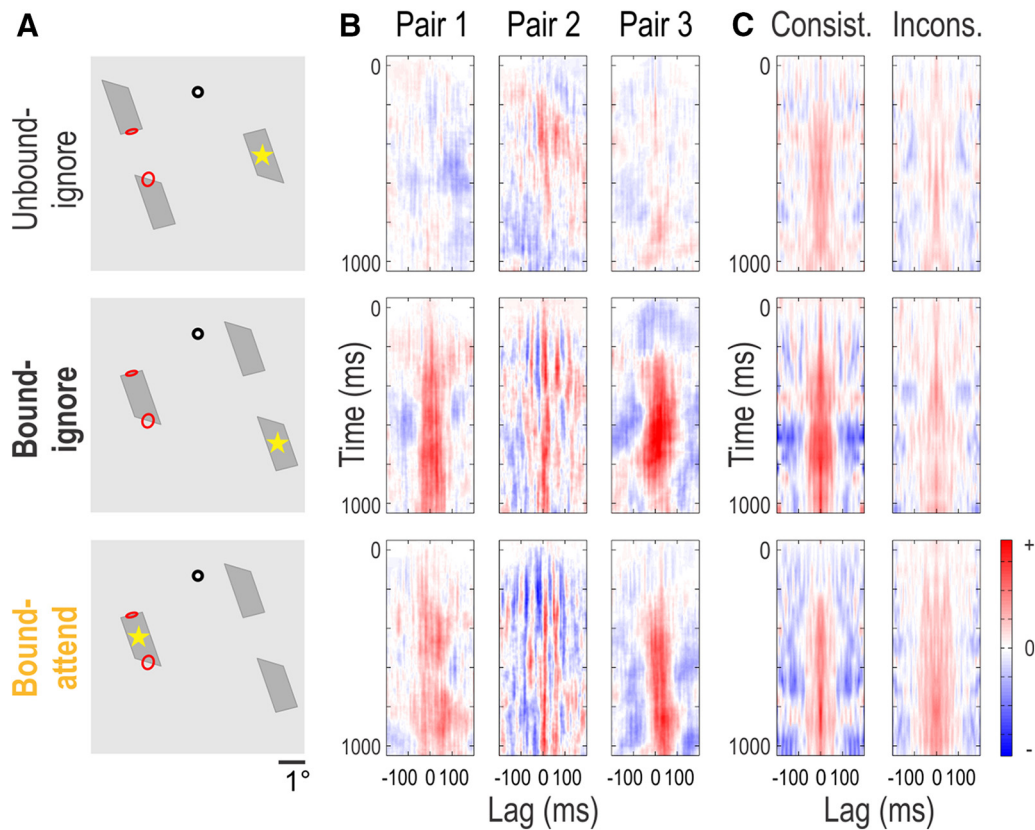


Figure 3. Effect of object configuration and selective attention on spike coincidences over the course of the trial. **A**, Example set of display configurations (as for neuron pair 1) with fixation spot (circle). Receptive fields (ellipses) and star (indicating attended figure) are not displayed to the animal. **B**, Frequency of spike coincidences over the course of the trial (250 ms sliding window) for three example pairs with consistent border-ownership preferences. Vertical axis, Time relative to figure onset; 1000 represents figure movement (cue to respond). Horizontal axis, Time lag between spikes of one neuron relative to the other. Color indicates coincidence counts relative to the level expected by chance; red, positive; blue, negative. Top, middle, and bottom panels correspond to conditions illustrated in **A**. **C**, As in **B** but for the populations of consistent pairs ($n = 38$) and inconsistent pairs ($n = 66$). Color scale bar indicates ± 300 coincidences/ s^2 for neuron pairs 1 and 2, ± 600 coincidences/ s^2 for neuron pair 3, and ± 80 coincidences/ s^2 for population means.

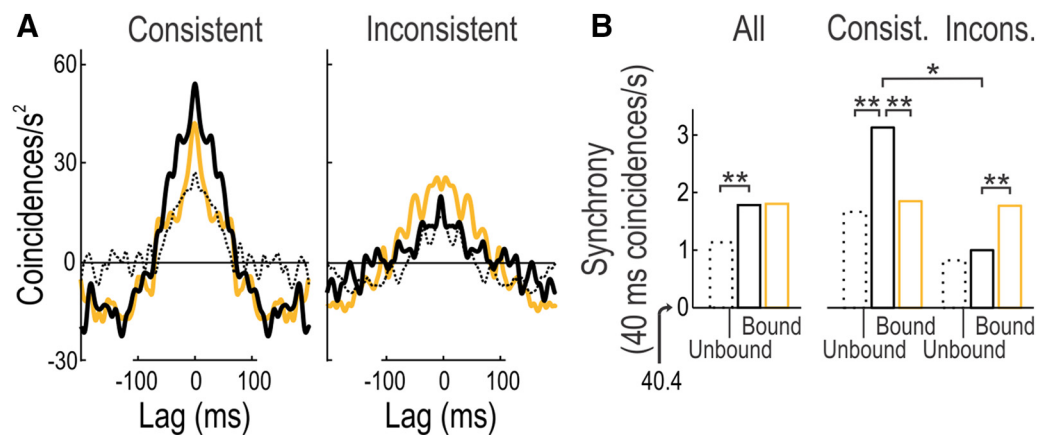


Figure 4. Synchrony in consistent and inconsistent pairs. **A**, Mean covariograms of the activity from 400 to 1050 ms after stimulus onset for consistent and inconsistent pairs in three conditions: unbound-ignore (dashed black line), bound-ignore (solid black line), and bound-attend (yellow line). The vertical axis indicates the increase in spike pairings beyond chance level as a function of time lag between spikes. **B**, Frequency of spike coincidences within 40 ms for all pairs pooled and for the consistent and inconsistent pairs separately. Line conventions as in **A**. The number at the curved arrow shows the frequency of coincidences expected by chance (average across groups and conditions). Note increase of synchrony with binding in consistent pairs, and low level of synchrony in inconsistent pairs. Significance, * $p < 0.05$, ** $p < 0.01$, randomization test.

To compare the different hypotheses, we calculated the rates of synchrony by counting spike coincidences within 40 ms in the three conditions for all pairs as well as for the two groups separately (Fig. 4B). We chose 40 ms because this was approximately the half-width of Gaussians fitted to the peaks in Figure 4A. When averaging over neuron pairs indiscriminately (All), we see a small

but significant increase of synchrony with binding (0.6 Hz, $p < 0.01$, randomization test used here and for all following tests unless stated otherwise), but no effect of attention ($p = 0.52$). Thus, there is no evidence in our data for a role of synchrony in coding attention. Regarding coding of binding-by-synchrony, the increase in frequency of coincidences of 0.6 Hz must be com-

pared with the frequency of coincidences expected by chance, which was ~ 40 Hz, and its random variation, which was ± 30 Hz (SD), as determined by integrating the cross-correlation function of the peristimulus time histograms over the interval ± 40 ms (pair types and conditions averaged). We are able to detect the small increase because we can analyze many responses, but it is not clear how the brain could distinguish in a single response these occasional extra coincidences from the large variations in the number of chance coincidences.

What remains hidden when pooling all pairs of neurons is that the synchrony differs strongly between consistent and inconsistent pairs (Fig. 4B). Critically, the mean excess synchrony in the bound-ignores condition is ~ 3 times higher in the consistent pairs than in the other pairs ($p < 0.05$), and binding doubles synchrony in the consistent pairs ($p < 0.01$), but has no effect in the other pairs ($p = 0.3$). Thus, binding selectively increases synchrony in consistent pairs ($p < 0.05$). These results do not depend on the exact choice of the integration window used for defining synchrony: the increase in frequency of coincidences with binding in the consistent pairs was significant with integration windows from 5 to 120 ms ($p < 0.05$). The difference between the consistent and inconsistent groups was also significant over this wide range ($p < 0.05$).

Possible effects of eye movements

We considered eye movements a possible source of variation of synchrony. We analyzed the influence of the experimental factors on (1) the variation of gaze position across trials and (2) the frequency of microsaccades.

The gaze position was not influenced by the experimental factors binding, attend and contrast polarity. The effects on mean gaze position during the static stimulus period were $< 0.011^\circ$ in every monkey (ANOVA; number of trials, 9,552; 10,710; and 5,130, respectively). The effect of such small deviations on the neuronal responses is negligible.

The frequency of microsaccades is of particular interest here because microsaccades can produce significant brief modulations of firing rate (Martinez-Conde and Macknik, 2008) that would be synchronized between neurons. Obviously, such modulations could not explain the observed difference in synchrony between consistent and inconsistent pairs of neurons, because eye movements would affect both groups of pairs equally. However, the experimental conditions might have influenced the frequency of microsaccades, and thus synchrony. We determined the frequency of microsaccades during the fixation period of each trial and analyzed the influence of the experimental factors binding, attend, and contrast polarity by ANOVA. As in our analysis of synchrony, we performed two analyses: one for the ignore

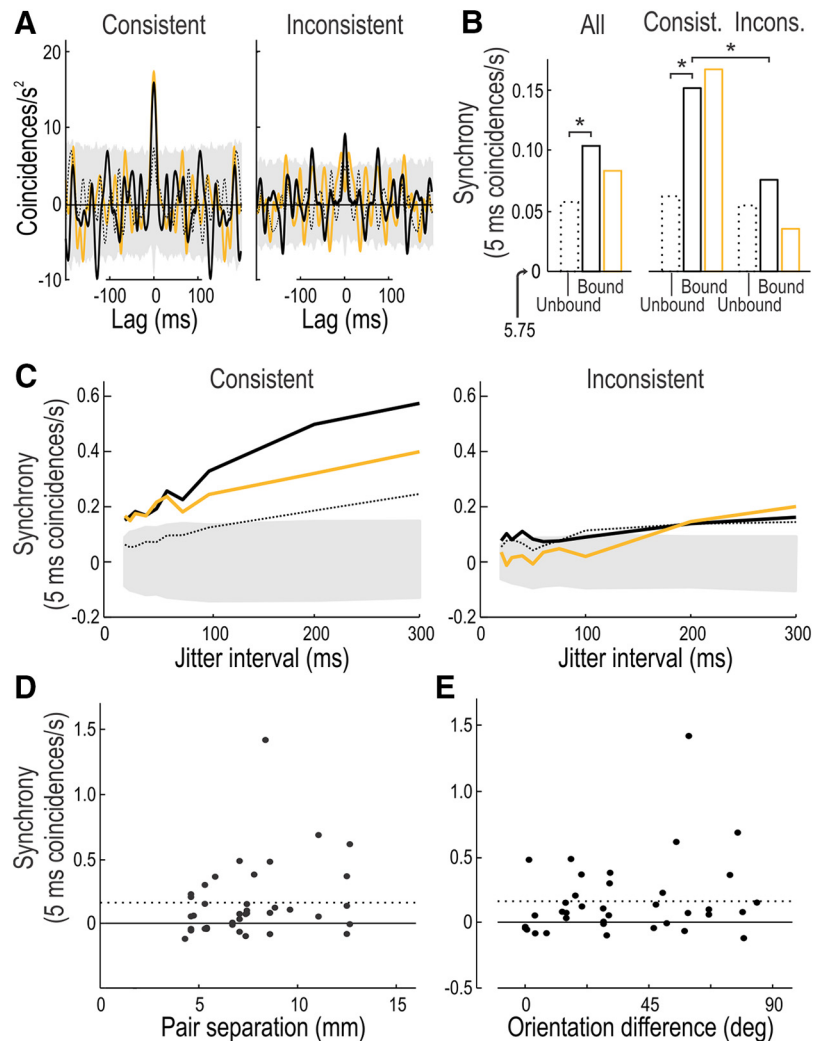


Figure 5. Precise synchrony. **A**, Mean covariograms only showing correlations with spike timing precision better than 20 ms. Conditions and line conventions as in Figure 4. Shading indicates, for the bound conditions, the 95% confidence limits of the null hypothesis that precise spike timing within 20 ms is irrelevant. **B**, Frequency of excess coincidences within 5 ms. The number at the curved arrow is the frequency of coincidences expected by chance (average across groups and conditions). Significance, $*p < 0.05$, randomization test. **C**, Frequency of excess coincidences (as in **B**) using different jitter intervals. Lines and shading as in **A**. **D**, Frequency of excess coincidences as a function of distance between neurons, average of bound conditions. Dashed line, Mean. **E**, Frequency of excess coincidences as a function of difference in preferred orientations between neurons. Dashed line, Mean. Note that millisecond synchrony was found even between neurons that were widely separated in cortex.

condition (attend = 0) with binding and contrast polarity as factors, and one for the bound condition (binding = 1) with attend and contrast polarity as factors. In the first analysis, we found small significant main effects of binding in the ignore trials (attend = 0) in each of the three monkeys. The effects were negative (-6.4% in BE; -3.6% in DW; and -3.0% in GW), i.e., the frequency of microsaccades was lower in the bound than in the unbound condition, which is opposite to the effects of binding on synchrony shown in Figure 4B ($+88\%$; the reason why the configuration influenced the frequency of microsaccades is not clear; however, the influence is very small). There were no significant effects of the factor contrast polarity. Meanwhile, the second analysis did not show any significant effects. Thus the attention condition did not influence the frequency of microsaccades. Note that the difference between the attention conditions is whether the attended figure was the one that stimulated the recorded neurons or one of the other figures, a factor that should not

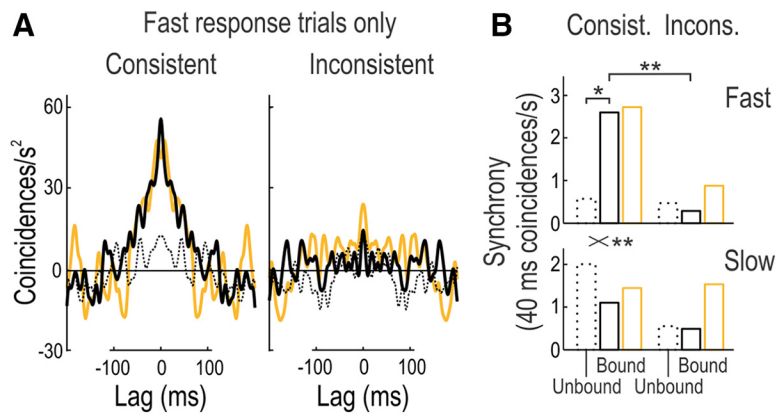


Figure 6. Correlation between synchrony and speed of behavioral response. **A**, Mean covariograms as in Figure 4, but showing the results for the quartile of trials with fastest reaction times. **B**, Synchrony in the fastest and slowest quartiles. Significance, * $p < 0.05$, ** $p < 0.01$, randomization test. Note enhanced effect of binding and enhanced differentiation between groups in fast-response trials compared with average trials (Fig. 4).

influence the frequency of microsaccades. In conclusion, microsaccades cannot explain the observed changes in synchrony.

Precision of coincidences

The fact that the covariograms of consistent pairs peak at zero lag suggests that synchrony is produced by common input, as postulated by the feedback hypothesis, rather than mutual interaction, as in a lateral propagation scheme. Given the large distances between our recording electrodes (3–14 mm), lateral propagation would produce signal delays due to the slow conduction of horizontal fibers (approximately one-tenth of the speed of white matter fibers; Girard et al., 2001). Such delays would produce higher correlations at non-zero lag times. Nevertheless, because of the relatively broad peaks in the mean covariograms of Figure 4A, one might argue that they could be the result of overlay of individual covariograms, some of which peaked at non-zero lag times.

To test for synchrony more rigorously, we used a perturbation method to extract synchrony that implies spike timing precision within a narrow interval (Amarasingham et al., 2012). For a given time interval δ , the method tests the null hypothesis that the exact spike timing within δ is irrelevant. From the experimental spike trains, surrogate spike trains are constructed that have the same variation of firing rate when sampled with time bins of length δ (the “jitter interval”), but within each bin the spike times are randomized with uniform distribution. For each pair of neurons, corresponding pairs of surrogate spike trains are generated and their cross-correlation functions are computed. Under the null hypothesis, the surrogate cross-correlation functions would statistically equal the experimental cross-correlation functions.

We calculated the difference between the experimental cross-correlation functions and the corresponding mean surrogate cross-correlation functions using $\delta = 20$ ms. The resulting new covariograms again showed significant peaks at zero lag for consistent pairs in the binding condition (ignore as well as attend), but not in the unbound condition, and not in the inconsistent pairs (Fig. 5A). We calculated the rates of synchrony by counting spike coincidences within 5 ms (the approximate half-width of Gaussians fitted to the peaks). This “tight synchrony” in consistent pairs was again significantly higher with binding (Fig. 5B, $p < 0.02$), and higher than in the inconsistent pairs ($p < 0.04$).

The effects of binding on synchrony and the differences between consistent and inconsistent pairs were not dependent on

the particular width δ of the jitter interval (Fig. 5C). Consistent pairs showed significant synchrony with binding even for intervals as narrow as 10 ms. This is strong evidence for common input through fast-conducting white matter fibers as a source of synchrony. In the millisecond range, synchrony in consistent pairs was similar for attended and ignored figures, and higher in both cases than in the unbound condition. When the jitter interval was widened (≥ 100 ms), the amount of synchrony increased ($p < 0.003$, Wilcoxon signed-rank test), and more so for ignored (Fig. 5C, solid black) than attended objects (yellow; $p < 0.01$ for comparison of attention effects at 10 vs 100 ms, Wilcoxon signed-rank test). Thus, attention affected only broad synchrony, as caused by relatively slow rate covariation. Attention

also reduced the trial-by-trial correlation between the mean firing rates (“noise correlation”; Cohen and Maunsell, 2009; Mitchell et al., 2009) in consistent pairs compared with inconsistent pairs (effect of attention on Pearson’s $r = -0.05$ vs $+0.06$, respectively, difference significant at $p < 0.03$).

Synchronous firing in widely separated neurons

Experiments in anesthetized animals (where feedback from higher levels is presumably reduced or absent) have shown that such tight synchrony falls off rapidly with distance between neurons, reaching zero at 4 mm (Smith and Kohn, 2008; for LFP and MUA coherence, see Frien and Eckhorn, 2000), which is approximately the maximum length of horizontal fibers in V1. Synchrony was also limited to pairs of neurons with similar orientation preferences (Kohn and Smith, 2005; Smith and Kohn, 2008). The grouping hypothesis predicts the opposite: to be flexible, the grouping mechanisms must encompass neurons with widely separated receptive fields and a variety of orientation preferences. Thus, we should see synchronous firing even between neurons with disparate receptive fields if they have consistent border-ownership selectivity and are activated by a common object.

We found that the jitter-reduced covariograms of the consistent pairs showed peaks of synchrony at zero lag even across large distances. The distribution of the strength of synchrony as a function of the distance between neurons shows no indication of a decline with distance (Fig. 5D; $p = 0.87$, $N = 38$, linear regression, one-sided test), and tight synchrony appeared in pairs separated by ≤ 13 mm. The occurrence of synchrony also did not depend on similarity of preferred orientations (Fig. 5E; $p = 0.87$, $N = 38$, linear regression, one-sided test).

Behavioral relevance

The grouping hypothesis predicts that behavioral responses will be facilitated by the feedback because it allows object-based modulation of feature signals: enhancement of the target and suppression of distracters. The differential modulation will improve the signal-to-noise ratio in downstream processing centers and speed up responses. Thus, if the strength of grouping feedback fluctuates from trial to trial, the grouping hypothesis predicts that stronger synchrony in the static presentation period of a trial should be followed by a faster behavioral response. As predicted, reaction time was negatively correlated with synchrony in consis-

tent pairs in the binding condition ($p < 0.02$, linear regression test). That is, reaction times were shorter as synchrony increased, with an improvement of -8.3 ms between the low and high synchrony quartiles. On the other hand, the inconsistent pairs showed no correlation between reaction times and synchrony ($p = 0.8$, linear regression test). To illustrate the difference in synchrony preceding fast and slow responses, we divided the trials into quartiles according to reaction time (Fig. 6). In the fast-response trials there was a significant increase of synchrony with binding in the consistent pairs ($p < 0.05$), whereas the slow trials showed no increase with binding ($p = 0.2$). The interaction between binding and response speed was significant ($p < 0.004$). The synchrony in the inconsistent pairs was not affected by binding in fast or slow trials ($p = 0.56$ and $p = 0.55$, respectively). In summary, stronger grouping is correlated with faster responses.

Coherent oscillations

Coherence (the covariance spectrum divided by the square root of the product of the power spectra) is a spectral measure of covariance between signals that has been used to detect rhythmic covariations between neural signals. Fries et al. (2001) showed that gamma-band coherence between LFPs in V4 increased with attention, but were not able to demonstrate a similar increase in the coherence spectra of spike trains. As they showed later (Fries et al., 2008), the reason for this was that coherence between spike trains is hard to detect due to the inherent high noise level in the coherence of spike trains, but this problem can be overcome using a more sensitive method that enhances the signal-to-noise ratio at the expense of spectral resolution.

We adopted this method to determine the effects of binding and attention on the coherence spectra of the consistent and inconsistent pairs of neurons. In contrast to the highly specific variation of spike synchrony, we found no evidence for fast oscillations in the coherence spectra (Fig. 7A). Neither binding nor attention increased spike coherence in the gamma range (binding $p = 0.6$, attend $p = 0.8$). Spike coherence differed only at low frequencies (Fig. 7B). In the consistent pairs, binding increased broadband low-frequency coherence for ignored objects ($p < 0.004$), but attention did not further increase coherence ($p = 0.5$). In the inconsistent pairs, binding did not affect coherence ($p = 0.4$), while attention increased the low-frequency coherence ($p < 0.004$). These observations at low frequencies are qualitatively consistent with the corresponding changes in shape of the covariograms (Fig. 4). We also computed coherence spectra between spikes on one electrode and the LCP on the other, but found no differences in low frequencies or the gamma range (Fig. 7C,D).

Discussion

The observation of elevated synchrony in neuron pairs with consistent border-ownership preferences and the finding that synchrony increased with binding in these pairs, but not in the others, confirm the main predictions of the grouping hypothesis: only pairs that share common feedback signals should become syn-

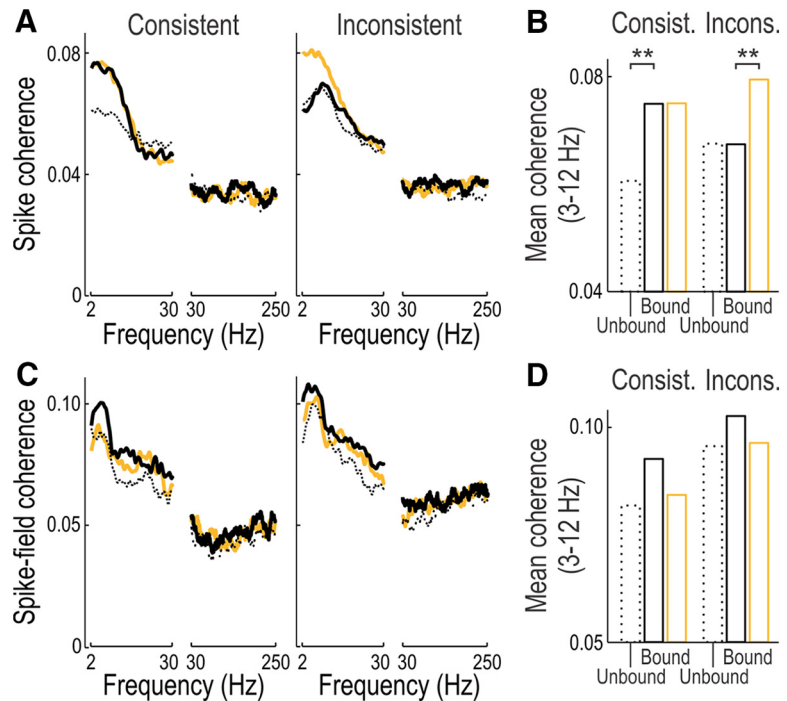


Figure 7. Coherence spectra. **A**, Mean spike coherence in the period from 400 to 1050 ms after stimulus onset. Conditions and line conventions as in Figure 4. **B**, Spike coherence energy in the frequency range 3–12 Hz. Significance, $**p < 0.01$, randomization test. **C, D**, As in **A** and **B** but showing coherence between spikes on one electrode and LCP on the other. Note absence of effects of binding and attention on coherence in the high-frequency range (30–250 Hz).

chronized (Fig. 1). Finding synchrony in a specific set of neurons as predicted is strong evidence for the proposed proto-object circuits.

Synchrony by modulatory feedback

Modulation by feedback was previously proposed to explain the short latency of border-ownership signals that emerge within 10–35 ms after the onset of the responses (Zhou et al., 2000; Sugihara et al., 2011). Given the low conduction velocity of intracortical fibers, it is hard to explain the speed of this extensive context integration by lateral propagation of signals in V2 or V1 (Craft et al., 2007; Zhang and von der Heydt, 2010). It was therefore suggested that border-ownership-selective neurons receive facilitating feedback from a higher level through fast-conducting white matter fibers (Zhou et al., 2000). Our finding of synchrony between consistent pairs of border-ownership neurons is direct evidence for such feedback. We propose that feedback grouping circuits explain not only the short latency of border-ownership signals, but also are the key to understanding the mechanism of proto-objects.

The shape of the consistent pairs' covariograms (Fig. 4A) is the signature of a specific mechanism. Previous studies of border-ownership selectivity invariably showed that the context effect is modulatory; context elements alone, without a stimulus in the classical receptive field, do not evoke spikes (Zhou et al., 2000; Zhang and von der Heydt, 2010). Theories and experimental studies suggest that NMDA receptors are responsible for some of the modulatory effects of recurrent projections, while AMPA receptors carry the feedforward transmission of visual signals (Johnson and Burkhalter, 1994; Lumer et al., 1997; Self et al., 2012). Whereas depolarizing synapses (e.g., AMPA) can trigger synchronous spikes in target neurons, common input through modulatory synapses will only amplify the synaptic currents originating from AMPA receptors. Thus, if grouping-cell feedback is mediated by NMDA receptors, then that input would not be

expected to directly trigger spikes in the target neurons. Instead, the grouping feedback would induce a simultaneous enhancement of the visually evoked responses that lasts for a duration given by the time constant of the NMDA synapses. The relatively long time constant of these synapses matches the width of the consistent pairs' covariograms (for the bound-ignore condition, for example, fitting a Gaussian returns $\sigma = 46$ ms, $r^2 = 0.97$). The sharp peak component of the covariograms revealed by the jitter method (Fig. 5A) is consistent with the time course of NMDA-mediated synaptic currents, which rise within a few milliseconds, in contrast to their slow decay (Hestrin et al., 1990). Thus the sharp peaks at zero lag time are strong evidence for common grouping-cell input.

Synchrony between neurons may not be associated with binding or attention in general, but only when this sort of grouping circuit is active. Several recent studies found no evidence for a role of synchrony in feature binding or attention coding in area V1 and the middle temporal area (Lamme and Spekreijse, 1998; but see Woelber et al., 2002; Roelfsema et al., 2004; Palanca and DeAngelis, 2005; Chen et al., 2014), suggesting that similar grouping circuits may not have been recruited. However, it should be noted that those and many other studies analyzed synchrony in multiunit activity (Lamme and Spekreijse, 1998; Fries et al., 2001; Woelber et al., 2002; Roelfsema et al., 2004; Palanca and DeAngelis, 2005) or LFPs (Fries et al., 2001; Woelber et al., 2002; Palanca and DeAngelis, 2005), whereas our results are based on single-cell recordings. By taking into account the specific role of the neurons in processing, our experiment revealed synchrony that might not have been visible otherwise. Only after characterizing the selectivity of the individual neurons for orientation and border ownership were we able to distinguish the different types of pairs and examine the specific predictions of the grouping hypothesis.

Synchrony as a coding mechanism?

While coding-by-synchrony theories postulate synchrony between neurons indiscriminately, the grouping model (Fig. 1) explains the observed specificity of synchrony. Only those neurons whose border-ownership preferences are consistent with a given object participate in its representation and thus exhibit synchrony. Although binding did increase the mean synchrony in our experiments, it added only $\sim 1\%$ to the average frequency of coincidences. It is hard to see how downstream centers could distinguish the few additional "meaningful" coincidences from random fluctuations in the large number of chance coincidences (40 ± 30 Hz; i.e., fluctuations of 75%). Thus, our study agrees with previous studies (Lamme and Spekreijse, 1998; Roelfsema et al., 2004; Palanca and DeAngelis, 2005) in concluding that synchrony does not provide a robust binding signal. In our grouping theory, the amount of synchrony is not critical because it is not postulated as a code. Here synchrony reflects a specific neural connectivity, and it is this connectivity that underlies the enhancement of feature responses with binding and produces additional enhancement under selective attention (Craft et al., 2007; Qiu et al., 2007; Mihalas et al., 2011).

Synchrony and attention

Our results show that the formation of proto-objects does not require attention. Synchrony also emerged for objects that the subjects tried to ignore, and attention actually reduced synchrony (Fig. 4B). This is in contrast to theories that emphasized the role of attention in feature binding (Treisman and Gelade, 1980). We argue that binding consists of the activation of reflexive grouping circuits (Fig. 1) that provide the structure for signal enhancement

in selective attention as observed in the visual cortex (Roelfsema et al., 1998; Qiu et al., 2007; Wannig et al., 2011). This concept also helps to explain the reorganization and persistence of figure-ground signals that parallel object perception (Khayat et al., 2004; Qiu and von der Heydt, 2007; O'Herron and von der Heydt, 2009, 2011) and the remapping of these signals across saccades and object movements (O'Herron and von der Heydt, 2013). For example, O'Herron and von der Heydt (2009) showed that border-ownership signals in V2 persist after the figure-ground cues are removed from the display, and that persisting signals can be found even after interruption of the V2 activity, implying persistence of signals outside V2; that is, at the level of grouping cells.

The attention effect needs further explanation. While enhancing firing rates, as predicted, attention decreased synchrony (Fig. 4). If the enhancement is produced by the grouping mechanism (Fig. 1) synchrony should also increase. However, it is important to remember that the reduction of synchrony in our results is relative to the ignore condition in which attention is focused on another object. Attention may have a dual role: suppression of distractors as well as enhancement of the target. Suppression of distractors, which is not included in the model of Figure 1, might also involve grouping cells, but with additional inhibitory interneurons. It is conceivable that fluctuations of grouping-cell activity in the suppression mechanism produce higher synchrony at the ignored objects than the enhancement of signals at the target (cf. Chen et al., 2014). The observation that the covariogram for the ignore condition was broader than that for the attend condition (Fig. 4) and the fact that the reduction of synchrony by attention was not apparent in the tight synchrony (Fig. 5A–C) would be consistent with the assumption of a separate suppressive mechanism that is more sluggish than the enhancement mechanism.

According to our theory, salient objects are represented by discrete peaks of activity in the grouping-cell layer. Finding elevated synchrony with binding even for ignored objects indicates that grouping cells enhance by feedback the contour signals of all the objects. Thus, the pattern of grouping-cell activity forms a preliminary cortical "object map." The top-down attention mechanism can select from this map simply by boosting the grouping cells corresponding to the object of interest relative to the others. Because each grouping cell connects to a large number of feature neurons, controlling their gain, the top-down attention mechanism can select a myriad of widely distributed feature signals simply by activating a small number of grouping cells (Mihalas et al., 2011). Thus, the peaks of activity in the grouping-cell layer might correspond to the proto-objects of perception (Rensink, 2000).

In support of this view we have shown that the proto-object map covaries with the performance in the shape discrimination task used in this study. The monkeys' fastest responses were preceded by a clearer map, as indicated by strong synchrony in consistent pairs of neurons in the binding condition and low synchrony in the other pairs (Fig. 6). We speculate that the proto-object map might provide the structure for object-based selective attention in general.

References

- Amarasingham A, Harrison MT, Hatsopoulos NG, Geman S (2012) Conditional modeling and the jitter method of spike resampling. *J Neurophysiol* 107:517–531. [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](#)
- Bartlett MS (1936) The square root transformation in the analysis of variance. *J R Stat Soc Suppl* 3:68–78. [CrossRef](#)
- Baylis GC, Driver J (2001) Shape-coding in IT cells generalizes over contrast

- and mirror reversal, but not figure-ground reversal. *Nat Neurosci* 4:937–942. [CrossRef Medline](#)
- Bokil H, Purpura K, Schoffelen JM, Thomson D, Mitra P (2007) Comparing spectra and coherences for groups of unequal size. *J Neurosci Methods* 159:337–345. [CrossRef Medline](#)
- Brody CD (1999) Disambiguating different covariation types. *Neural Comput* 11:1527–1535. [CrossRef 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](#)
- Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12:1594–1600. [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](#)
- Duncan RO, Albright TD, Stoner GR (2000) Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context. *J Neurosci* 20:5885–5897. [Medline](#)
- Frien A, Eckhorn R (2000) Functional coupling shows stronger stimulus dependency for fast oscillations than for low-frequency components in striate cortex of awake monkey. *Eur J Neurosci* 12:1466–1478. [CrossRef Medline](#)
- Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563. [CrossRef Medline](#)
- Fries P, Womelsdorf T, Oostenveld R, Desimone R (2008) The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci* 28:4823–4835. [CrossRef Medline](#)
- Girard P, Hupé JM, Bullier J (2001) Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* 85:1328–1331. [Medline](#)
- Gray CM, König P, Engel AK, Singer W (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338:334–337. [CrossRef Medline](#)
- Hestrin S, Sah P, Nicoll RA (1990) Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron* 5:247–253. [CrossRef Medline](#)
- Huang X, Albright TD, Stoner GR (2008) Stimulus dependency and mechanisms of surround modulation in cortical area MT. *J Neurosci* 28:13889–13906. [CrossRef Medline](#)
- Jarvis MR, Mitra PP (2001) Sampling properties of the spectrum and coherence of sequences of action potentials. *Neural Comput* 13:717–749. [CrossRef Medline](#)
- Johnson RR, Burkhalter A (1994) Evidence for excitatory amino acid neurotransmitters in forward and feedback corticocortical pathways within rat visual cortex. *Eur J Neurosci* 6:272–286. [CrossRef Medline](#)
- Kahneman D, Treisman A, Gibbs BJ (1992) The reviewing of object files: object-specific integration of information. *Cogn Psychol* 24:175–219. [CrossRef Medline](#)
- Kanizsa G (1979) Organization in vision. Essays on Gestalt perception. New York: Praeger.
- Khayat PS, Spekreijse H, Roelfsema PR (2004) Correlates of transsaccadic integration in the primary visual cortex of the monkey. *Proc Natl Acad Sci U S A* 101:12712–12717. [CrossRef Medline](#)
- Koffka K (1935) Principles of Gestalt psychology. New York: Harcourt, Brace and World.
- Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci* 25:3661–3673. [CrossRef Medline](#)
- Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15:1605–1615. [Medline](#)
- Lamme VA, Spekreijse H (1998) Neuronal synchrony does not represent texture segregation. *Nature* 396:362–366. [CrossRef Medline](#)
- Li W, Piéch V, Gilbert CD (2006) Contour saliency in primary visual cortex. *Neuron* 50:951–962. [CrossRef Medline](#)
- Lumer ED, Edelman GM, Tononi G (1997) Neural dynamics in a model of the thalamocortical system. I. Layers, loops and the emergence of fast synchronous rhythms. *Cereb Cortex* 7:207–227. [CrossRef Medline](#)
- Martinez-Conde S, Macknik SL (2008) Fixational eye movements across vertebrates: comparative dynamics, physiology, and perception. *J Vis* 8(14):28 1–16. [CrossRef Medline](#)
- Mihalas S, Dong Y, von der Heydt R, Niebur E (2011) Mechanisms of perceptual organization provide auto-zoom and auto-localization for attention to objects. *Proc Natl Acad Sci U S A* 108:7583–7588. [CrossRef Medline](#)
- Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63:879–888. [CrossRef Medline](#)
- Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. *Biophys J* 76:691–708. [CrossRef Medline](#)
- Nakayama K, Shimojo S, Silverman GH (1989) Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18:55–68. [CrossRef Medline](#)
- Nakayama K, He ZJ, Shimojo S (1995) Visual surface representation: a critical link between lower-level and higher-level vision. In: *Invitation to cognitive science* (Kosslyn SM, Osherson DN, eds), pp 1–70. Cambridge, MA: MIT.
- Niebur E, Koch C (1994) A model for the neuronal implementation of selective visual attention based on temporal correlation among neurons. *J Comput Neurosci* 1:141–158. [CrossRef Medline](#)
- O’Herron P, von der Heydt R (2009) Short-term memory for figure-ground organization in the visual cortex. *Neuron* 61:801–809. [CrossRef Medline](#)
- O’Herron PJ, von der Heydt R (2011) Representation of object continuity in the visual cortex. *J Vis* 11(2):pii: 12. [CrossRef Medline](#)
- O’Herron P, von der Heydt R (2013) Remapping of border ownership in the visual cortex. *J Neurosci* 33:1964–1974. [CrossRef Medline](#)
- Palanca BJ, DeAngelis GC (2005) Does neuronal synchrony underlie visual feature grouping? *Neuron* 46:333–346. [CrossRef Medline](#)
- Pylyshyn ZW (2001) Visual indexes, preconceptual objects, and situated vision. *Cognition* 80:127–158. [CrossRef Medline](#)
- 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](#)
- Qiu FT, Sugihara T, von der Heydt R (2007) Figure-ground mechanisms provide structure for selective attention. *Nat Neurosci* 10:1492–1499. [CrossRef Medline](#)
- Rensink RA (2000) Seeing, sensing, and scrutinizing. *Vision Res* 40:1469–1487. [CrossRef Medline](#)
- Rensink RA, Enns JT (1998) Early completion of occluded objects. *Vision Res* 38:2489–2505. [CrossRef Medline](#)
- Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. *Neuron* 26:703–714. [CrossRef Medline](#)
- Roelfsema PR, Lamme VA, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395:376–381. [CrossRef Medline](#)
- Roelfsema PR, Lamme VA, Spekreijse H (2004) Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nat Neurosci* 7:982–991. [CrossRef Medline](#)
- Self MW, Kooijmans RN, Supèr H, Lamme VA, Roelfsema PR (2012) Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proc Natl Acad Sci U S A* 109:11031–11036. [CrossRef Medline](#)
- Smith MA, Kohn A (2008) Spatial and temporal scales of neuronal correlation in primary visual cortex. *J Neurosci* 28:12591–12603. [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](#)
- Treisman AM, Gelade G (1980) A feature integration theory of attention. *Cogn Psychol* 12:97–136. [CrossRef Medline](#)
- Vogels R, Spileers W, Orban GA (1989) The response variability of striate cortical neurons in the behaving monkey. *Exp Brain Res* 77:432–436. [CrossRef Medline](#)
- von der Malsburg C (1986) Am I thinking assemblies? In: *Brain theory. Proceedings of the First Trieste Meeting on Brain Theory, October 1–4, 1984* (Palm G, Aertsen A, eds), pp 161–176. Berlin: Springer.
- Wannig A, Stanisor L, Roelfsema PR (2011) Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex. *Nat Neurosci* 14:1243–1244. [CrossRef Medline](#)
- Wolbert T, Eckhorn R, Frien A, Bauer R (2002) Perceptual grouping correlates with short synchronization in monkey prestriate cortex. *Neuroreport* 13:1881–1886. [CrossRef Medline](#)
- Zhang NR, von der Heydt R (2010) Analysis of the context integration mechanisms underlying figure-ground organization in the visual cortex. *J Neurosci* 30:6482–6496. [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](#)