Royal Netherlands Academy of Arts and Sciences (KNAW)
KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex.

Poort, J.; Self, M.W.; Van Vugt, B.; Malkki, H.A.; Roelfsema, P.R.

published in
Cerebral Cortex
2016

DOI (link to publisher)
10.1093/cercor/bhw235

document version
Publisher's PDF, also known as Version of record

Link to publication in KNAW Research Portal

citation for published version (APA)
Poort, J., Self, M. W., Van Vugt, B., Malkki, H. A., & Roelfsema, P. R. (2016). Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex. *Cerebral Cortex*, *26*, 3964-3976. https://doi.org/10.1093/cercor/bhw235

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
pure@knaw.nl

Download date: 17. Oct. 2026

Cerebral Cortex, October 2016;26: 3964–3976

doi: 10.1093/cercor/bhw235
Advance Access Publication Date: 13 August 2016
Original Article

ORIGINAL ARTICLE

Texture Segregation Causes Early Figure Enhancement and Later Ground Suppression in Areas V1 and V4 of Visual Cortex

Jasper Poort[1,2,3], Matthew W. Self[1], Bram van Vugt[1], Hemi Malkki[1], and Pieter R. Roelfsema[1,4,5]

[1]Department of Vision and Cognition, Netherlands Institute for Neuroscience, An Institute of the Royal Netherlands Academy of Arts and Sciences, Amsterdam, 1105 BA, the Netherlands, [2]Department of Neuroscience, Physiology and Pharmacology, University College London, London, WC1E 6DE, UK, [3]Sainsbury Wellcome Centre for Neural Circuits and Behaviour, University College London, London, W1T 4JG, UK, [4]Department of Integrative Neurophysiology, Centre for Neurogenomics and Cognitive Research, Vrije Universiteit, Amsterdam, 1081 HV, the Netherlands, and [5]Psychiatry Department, Academic Medical Center, Amsterdam, 1105 AZ, the Netherlands

J. Poort and M. W. Self have contributed equally to this work.
Address correspondence to J. Poort, Sainsbury Wellcome Centre for Neural Circuits and Behaviour, University College London, 25 Howland Street, London, W1T 4JG, UK. Email: j.poort@ucl.ac.uk.

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

Abstract

Segregation of images into figures and background is fundamental for visual perception. Cortical neurons respond more strongly to figural image elements than to background elements, but the mechanisms of figure–ground modulation (FGM) are only partially understood. It is unclear whether FGM in early and mid-level visual cortex is caused by an enhanced response to the figure, a suppressed response to the background, or both.

We studied neuronal activity in areas V1 and V4 in monkeys performing a texture segregation task. We compared texture-defined figures with homogeneous textures and found an early enhancement of the figure representation, and a later suppression of the background. Across neurons, the strength of figure enhancement was independent of the strength of background suppression.

We also examined activity in the different V1 layers. Both figure enhancement and ground suppression were strongest in superficial and deep layers and weaker in layer 4. The current–source density profiles suggested that figure enhancement was caused by stronger synaptic inputs in feedback-recipient layers 1, 2, and 5 and ground suppression by weaker inputs in these layers, suggesting an important role for feedback connections from higher level areas. These results provide new insights into the mechanisms for figure–ground organization.

Key words: cortical layers, figure–ground segmentation, perceptual organization, vision, visual cortex

© The Author 2016. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

The assignment of image elements to figure or background is an elementary step in visual perception. A powerful illustration of this process is the face–vase illusion (Fig. 1A), where our interpretation of the image alternates (Rubin 1915). The assignment of image regions to figure or ground has a profound influence on perception, because image elements that are part of figures receive preferential processing and leave stronger memory traces (Driver and Baylis 1996; Baylis and Cale 2001). The perceptual status of the ground regions is less clear. One study suggested that background features are not processed up to a perceptual level (Baylis and Cale 2001), but others suggested that background regions are actively suppressed (DeSchepper and Treisman 1996; Peterson and Skow 2008; Salvagio et al. 2012). This question can also be formulated at the level of neuronal processing: does figure–ground segregation enhance the neuronal representation of the figure, suppress the background, or both (Fig. 1D)?

Previous studies of neuronal activity in primary visual cortex (V1) and mid-level area V4 (Lamme 1995; Poort et al. 2012) during texture segregation found that neurons respond more vigorously to image elements of a figure than to elements of the background (Fig. 1B). This response difference is called figure–ground modulation (FGM). Interestingly, FGM is strongest in the superficial and deep layers of V1 and weakest in input layer 4, and it is associated with a pattern of synaptic activity that suggests an important role for feedback from higher visual areas (Self et al. 2013). However, the precise contributions of figure enhancement and ground suppression to FGM in the texture segregation task remain unknown. Strong suppressive effects were observed by Landman et al. (2003), who demonstrated that the activity elicited by background elements in V1 decreases with the number of figures present in a display. In contrast, one functional magnetic resonance imaging (fMRI) study has demonstrated that responses elicited by figures are enhanced in visual cortex (V1–V4) (Scholte et al. 2008). However, another fMRI study (Likova and Tyler 2008) did not find figure enhancement but only background suppression. Both studies also found effects of figure–ground perception in extra-striate cortex, but they could not resolve activity related to figure and background regions. In a related study on contour-grouping, Chen et al. (2014) examined activity in areas V1 and V4 of monkeys trained to perceive an elongated contour formed by collinear line elements among randomly oriented distractor elements. The representation of elements of the contour was enhanced in V1, and the activity elicited by the randomly oriented line elements was suppressed (see Gilad et al. 2013, for similar results using voltage-sensitive dye imaging

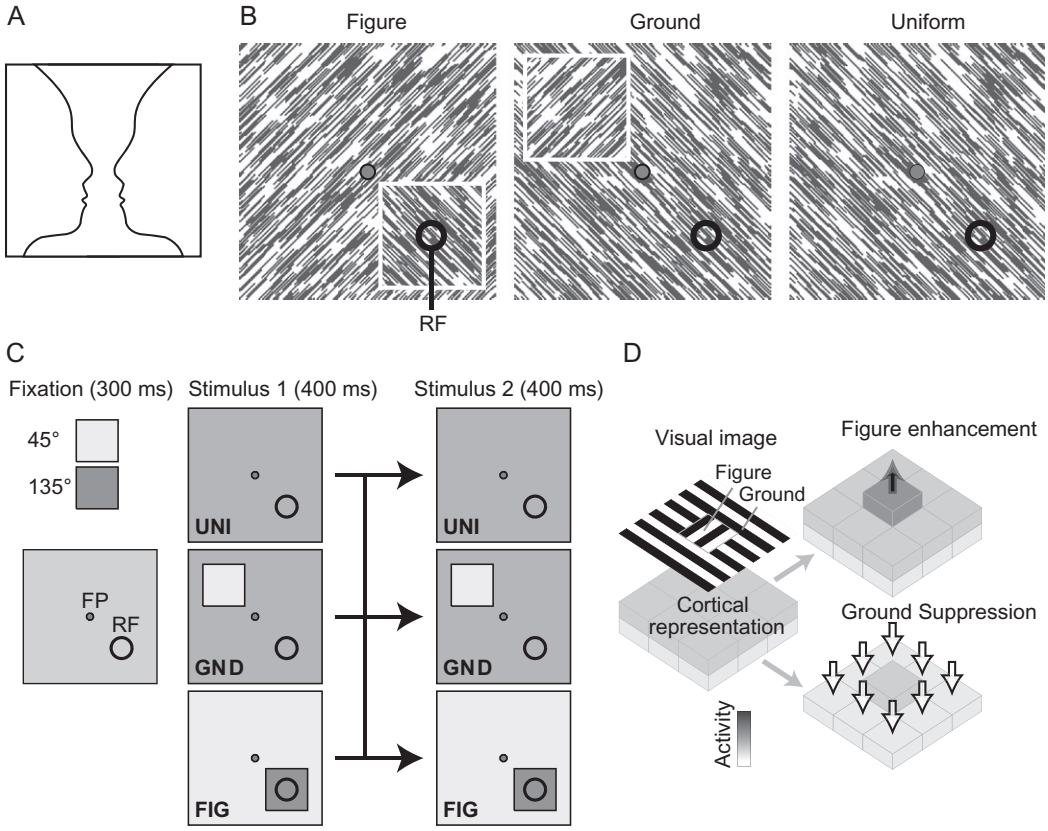

Figure 1. Illustration of figure–ground organization and the behavioral paradigm. (A) Face–vase illusion. This classic ambiguous figure–ground display causes perception to alternate between 2 faces or a vase on a formless background. (B) Illustration of the 3 types of stimuli. The black circular outline illustrates the position of a receptive field (RF), the central gray circle indicates the fixation point (FP). Left, RF is on a figure (white square box, not visible to the monkey); Middle, RF is on the background; Right, RF is on a uniform texture. Note that the image elements in the RF are the same in all conditions. (C) In Experiment 1, each trial started when the monkey directed gaze to the FP and kept fixation for an additional 300 ms. Fixation was followed by 2 successive stimuli. The first stimulus consisted of one of the 3 configurations in panel A and was presented for 400 ms. The second stimulus was again one of the 3 configurations, which was presented for an additional 400 ms. Next, the fixation point disappeared and the monkey was rewarded after making a saccade to the location of the figure, if one was present. If no figure was present (uniform texture), the monkey had to maintain fixation for an additional 250 ms, after which the animal was rewarded. (C) Different scenarios for figure-ground modulation (FGM). FGM can be the result of figure enhancement or ground suppression, or a combination of these 2 processes.

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

in V1). A recent fMRI study also reported an enhanced contour representation combined with a suppression of randomly oriented contours (Strother et al. 2012). Both studies (Strother et al. 2012; Chen et al. 2014) also revealed effects in extra-striate cortex, but again, contour and background responses could not be measured separately.

In the texture segregation task, the contributions of figure enhancement and ground suppression to texture segregation and their timing (Tsotsos et al. 2008) remain unclear, and the role of these 2 processes in extra-striate cortex is generally unknown. Furthermore, previous studies did not address the influence of activity enhancement and suppression in the different cortical layers. We therefore recorded from V1 and V4, a higher area that plays an important role in texture segregation (Merigan 1996; Allen et al. 2009), to address the following questions: 1) how do figure enhancement and ground suppression contribute to FGM in V1 and extra-striate cortex during texture segregation? 2) What is the profile of figure enhancement and ground suppression across the layers in area V1? 3) Do neurons with figure enhancement also exhibit ground suppression or do these 2 processes influence different neuronal circuits?

Materials and Methods

Visual stimulus and behavioral paradigm

We conducted 2 experiments. In Experiment 1, we investigated the contribution of figure enhancement and ground suppression in a texture segregation task in areas V1 and V4, and in Experiment 2, we investigated the laminar profile of suppression and enhancement in area V1. The general aim of the experiments was to measure figure enhancement and ground suppression. We isolated the contribution of figure enhancement by comparing activity elicited by a figure in the neurons' receptive field (RF) to that elicited by a homogeneous texture (Fig. 1B, left vs. right panel). We isolated ground suppression by comparing activity elicited by the homogeneous texture with that elicited by the ground condition in which there was a figure remote from the neurons' RF (Fig. 1B, right vs. middle panel).

Three monkeys participated in Experiments 1 and 2 monkeys in Experiment 2. They were seated at a distance of 0.75 m from a monitor (width: 0.4 m) with a resolution of 1024 × 768 pixels and a frame rate of 110 Hz (85 Hz in Experiment 2).

In Experiment 1, the visual stimulus consisted either of a square figure with oriented line elements (32 pixels long, 0.93°, and 2 pixels wide) on a background with an orthogonal orientation, or it consisted of a homogeneous texture (Fig. 1B). To construct the stimulus, we first made 4 full-screen base textures, 2 with an orientation of 45° and 2 with an orientation of 135°. A base texture was made by randomly placing 13 000 black line elements (luminance 2.8 cd m^{-2}) with a given orientation on a white (luminance 94 cd m^{-2}) background. We then created full-screen stimuli for the figure and ground conditions by copying a square 4° × 4° region of a 45° base texture onto a 135° base texture or by copying the same square region of a 135° base texture onto a 45° base texture. In the uniform condition, we presented only a 45° or 135° base texture that covered the full screen. To analyze the FGM, we averaged neuronal responses to the 2 complimentary stimuli, thereby ensuring that the RF was stimulated on average by the same set of local features, regardless of whether the RF was on the figure, background, or homogeneous texture (see Fig. 1B).

A trial started as soon as the monkey's eye position was within a 1° × 1° window centered on the fixation point (FP) (0.58°), presented on a gray background (luminance 34 cd m^{-2}). The monkey had to maintain fixation within the fixation window until cued to make a saccade by the disappearance of the FP. In Experiment 1, the monkey saw 2 figure–ground stimuli that were presented successively. When the monkey had kept fixation for 300 ms, the first stimulus was presented (period 1, 400 ms). It consisted of a figure in the RF (figure condition), a figure that was not in the RF (ground condition), or a uniform texture. The figure could appear at 1 of 4 locations: in 3 monkeys, we recorded data from 2 V1 electrode arrays (see below), so that RFs were clustered at 2 positions in the visual field. Therefore, we used 2 figure positions that were centered on one of the RF clusters and 2 corresponding positions at the same eccentricity as the RF clusters, rotated by 180°. After period 1, the second stimulus was presented that was again a figure, a ground, or uniform condition (period 2, 400 ms). After period 2, the FP disappeared, and the monkey had to make a saccade to the target window of 4° × 4° centered on the location of the figure to obtain a drop of apple juice as a reward (Fig. 1C). If there was no figure present in period 2 (uniform condition), the monkey was rewarded if he maintained fixation for an additional 250 ms. The monkeys detected figures with high accuracy (98% correct for monkey 1, 94% for monkey 2, and 96% for monkey 3). The accuracy was lower in catch trials without a figure (92% for monkey 1, 63% for monkey 2, and 74% for monkey 3) because the monkeys had to maintain fixation for a longer duration. We only included correct trials in all of our analyses.

In Experiment 2, there was only 1 epoch with a full-screen texture (5,345 line elements per texture with a width of 1 pixel and a length of 16 pixels). In 75% of trials, the texture contained a figure (4° × 4°). The figure was placed in the RF (figure condition), at one of 2 locations situated at the same eccentricity but at 120° away from the RF (ground condition). The animal had to maintain fixation for 300 ms, after which the fixation dot disappeared and the monkey had to make an eye movement to a 4° × 4° window centered on the figure. On the other 25% of trials, a uniform texture was presented and the animal was rewarded for maintaining fixation for an additional 400 ms after the fixation dot was extinguished. The performance in detecting figures was above 95% correct for both monkeys. The accuracy in catch trials was 77% for monkey 4 and 89% for monkey 5.

Surgical procedures

We used the same surgical protocol as described previously (Poort et al. 2012; Self et al. 2012). The monkeys underwent 2 surgeries under general anesthesia that was induced with ketamine (15 mg kg^{-1} injected intramuscularly) and maintained after intubation by ventilation with a mixture of 70% N_2O and 30% O_2, and supplemented with 0.8% isoflurane, fentanyl (0.005 mg kg^{-1} intravenously), and midazolam (0.5 mg kg^{-1} h^{-1} intravenously). In the first surgery, we implanted a head holder. In the second surgery, we implanted arrays of 4 × 5 electrodes (Cyberkinetics Neurotechnology Systems Inc.) in areas V1 and V4 for Experiment 1, and a chamber above V1 over a small craniotomy for the laminar recordings of Experiment 2. All procedures complied with the NIH Guide for Care and Use of Laboratory Animals (National Institutes of Health) and were approved by the Institutional Animal Care and Use Committee of the Royal Netherlands Academy of Arts and Sciences.

Recording of neuronal activity

In Experiment 1, we recorded multiunit activity in 2 monkeys that were chronically implanted with electrode arrays in V1

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

and V4, and 1 monkey with arrays only in V1. In Experiment 2, we recorded from V1 of 2 monkeys (they did not take part in Experiment 1) using a multicontact laminar probe ('U-probe', Plexon Inc.) that was inserted into V1, as described previously (Self et al. 2012). In both experiments, multiunit spiking activity (MUA) was recorded with a TDT (Tucker Davis Technologies) data acquisition system. As in previous studies (Legatt et al. 1980; Logothetis et al. 2001; Supèr and Roelfsema 2005), MUA signals were amplified, band-pass filtered (500–5000 Hz), full-wave rectified, and then low-pass filtered at 500 Hz and sampled at a rate of 763 Hz. The MUA signal contains spikes from neurons within ~150 μm of the electrode tip (Self et al. 2013), which corresponds to the distance over which a V1 cell can be recorded with single-unit recording. Accordingly, the MUA represents the pooled activity of a number of single units in the vicinity of the tip of the electrode, and the population response obtained with this method is therefore similar to the population response obtained by pooling across single units (Supèr and Roelfsema 2005; Cohen and Maunsell 2009). The eye position was measured with an eye tracker camera system (Thomas Recording) and sampled at a rate of 250 Hz.

For Experiment 2, we also computed the current–source density (CSD) from the local field potential (LFP). The LFP at each recording site was obtained by low-pass filtering the signal from the electrode below 200 Hz. The CSD was then calculated as:

$$CSD(x) = -\sigma \cdot \frac{\phi(x-h) - 2\phi(x) + \phi(x+h)}{h^2}$$

ϕ is the LFP voltage (in V), x is the point at which the CSD is calculated, h is the spacing of recording sites for the computation (here 200μm), and σ is tissue conductivity (we used 0.4 S m^{-1}). This equation yields the CSD in units of A m^{-3}, but we here report the CSD in physiologically more relevant units; μA mm^{-3}.

For each V1 recording site, we measured the RF by determining the onset and offset of the response to a slowly moving light bar in 8 movement directions (Kato et al. 1978). In Experiment 1, the median V1 RF area was 1.6 deg^2 (range 0.08–7.6 deg^2), and the median eccentricity was 4.02° (range 2.5°–6.9°). In Experiment 2, the median RF area was 2.2 deg^2 (range 0.39–15.8 deg^2) and the median eccentricity was 4.12° (range 1.8°–12°). In V4 (Experiment 1), we mapped RFs by presenting white dots (0.5 deg, luminance 82 cd m^{-2}) on a gray background (luminance 14 cd m^{-2}) at different positions of a grid (0.5 deg spacing). The hotspot of the V4 RF was defined as the position with the maximum response (median eccentricity 4.04°, range 0.79°–7.43°) and the RF borders as the locations where activity fell below 50% of the maximum (Motter 1993). Using this criterion, the median V4 RF area was 19.7 deg^2 (range 6.5–38 deg^2).

Data analysis

We quantified the visual responsiveness of neurons at each recording site by calculating the mean spontaneous activity level across all conditions *Sp* and the standard deviation *s* across trials in a 200 ms time window preceding stimulus onset. We then computed the peak response, *Pe*, by smoothing the average response across conditions with a moving window of 25 ms and taking the maximum during the stimulus period (0–300 ms after stimulus onset). The visual responsiveness index was computed as $VR = (Pe\text{-}Sp)/s$. Only recording sites with a good visual response ($VR > 3$) were included in the analyses. In Experiment 1, we included 102 V1 recording sites (40 in monkey 1, 33 in monkey 2, and 29 in monkey 3) and 36 in V4 (14 in monkey 1 and 22 in monkey 2). The number of recording sites in Experiment 2 will be specified below. MUA data from each recording site were normalized by subtracting *Sp* and subsequently dividing by (*Pe* –*Sp*).

FGM was computed as the difference between the responses evoked by the figure and background. To quantify the amount of figure enhancement, we computed the difference between the response evoked by the figure and the response evoked by the uniform texture (figure–uniform modulation, FUM). To quantify the amount of ground suppression, we computed the difference between the response elicited by the uniform texture and the background (with the figure at another location, outside the RF) (uniform–ground modulation, UGM).

We determined the latency of the visual responses, FUM and UGM by fitting a function $f(t)$ to the neural response (or response difference) (Thompson et al. 1996; Roelfsema et al. 2003). The function was derived from the assumptions that the onset of the response has a Gaussian distribution and that a fraction of the response dissipates exponentially which yields the following equation:

$$f(t) = d \cdot \exp(\mu\alpha + 0.5\sigma^2\alpha^2 - \alpha t) \cdot (G(t, \mu + \sigma^2\alpha, \sigma) + c \cdot G(t, \mu, \sigma),$$

where $G(t, \mu, \sigma)$ is a cumulative Gaussian density with mean μ and standard deviation σ, a^{-1} is the time constant of the dissipation, and c and d represent the contribution the non-dissipating and dissipating component (see Roelfsema et al. (2003) for details). We fitted $f(t)$ to the responses using the curve fitting toolbox in MATLAB (MathWorks) and defined the latency as the time point where the fitted function reached 33% of its maximum. To determine the significance of a latency difference between 2 conditions, we used a bootstrapping procedure. We randomly selected a number of recording sites equal to the original sample with replacement and fitted the latency in each condition and subtracted these latencies to obtain a distribution of the latency difference across 1000 repeats. The latency of the figure–uniform modulation was measured by fitting the curve to the difference between the response evoked by the figure and uniform texture, and the latency of the UGM by fitting a curve to the difference between the response evoked by the uniform texture and background.

For the latency analysis in the second stimulus period (Fig. 2E–H), we pooled across all conditions with a particular stimulus in the second period (figure, ground, or uniform), allowing the stimulus in the first period to vary (Fig. 1C). We ensured that the stimulus history was balanced so that the first stimulus did not predict the second stimulus. This enabled us to examine the relative timing of figure enhancement and ground suppression in the 2 stimulus periods. We note, however, that the transitions to the first and second stimulus differed. When we presented the first stimulus, the RF stimulation changed from a gray background to texture elements. In the second stimulus period, the RF stimulation changed from one texture to another. Thus, the transitions were not balanced, which may cause differences in the activity elicited in V1 and V4 by the 2 stimuli.

To quantify how reliably individual recording sites discriminated between the different stimulus conditions, we computed the *d*-prime: $d_{AB} = (m_A\text{–}m_B)/s$, where m_A and m_B are the mean responses in stimulus conditions *A* and *B*, and *s* is the pooled standard deviation. d_{FU} is a measure for the discrimination between a figure and a uniform texture, and d_{UG} is a measure for the discrimination between a uniform texture and the background. We quantified the correlation between the *d*-prime in

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

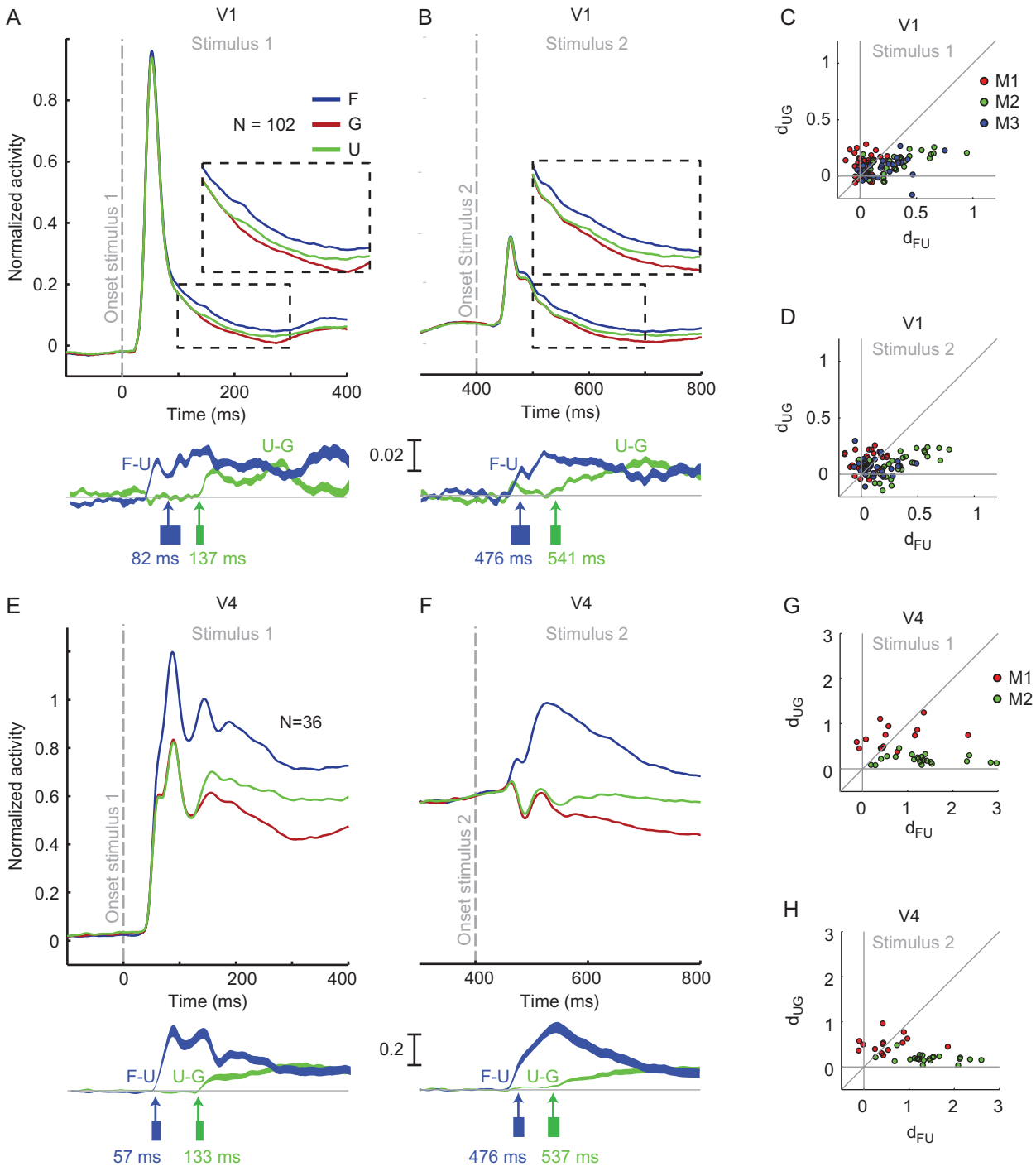

Figure 2. Figure enhancement and ground suppression in V1 and V4. (A) and (B) average neuronal activity in V1 in first (A) and second stimulus period (B) when the RF was on the figure (F, blue), background (G, red) or on the uniform texture (U, green). N = 102 recording sites. The lower panels illustrate the time course of figure enhancement (blue, figure minus homogeneous) and ground suppression (green, homogeneous minus ground). The arrows show the latency of figure enhancement and ground suppression and the bars denote the 95% CIs. (C) and (D) d-primes of figure enhancement (*x*-axis, d_{FU}) and ground suppression (*y*-axis, d_{UG}) for stimulus 1 (C) and stimulus 2 (D). Data from the 3 monkeys are shown in different colors. (E) and (F) activity in area V4. N = 36 recording sites. (G) and (H) d-primes in V4.

different conditions with Pearson's correlation coefficient, and used the Student's t distribution to assess significance.

Laminar analysis

In Experiment 2, we recorded from 30 penetrations in monkey 4 and 14 penetrations in monkey 5 with laminar electrodes with a spacing between neighboring electrodes of 100 µm. Part of the data of Experiment 2 have been used in a previous study (Self et al. 2013), but that study did not analyze the responses elicited by the homogeneous texture, which allowed us to separately determine the contribution of figure enhancement and ground suppression to FGM. We identified the depth of each recording site relative to the layer 4c/layer 5 boundary using

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

the CSD as described previously (Self et al. 2013). We then assigned each recording site to one of the 3 laminar compartments based on the distance of the recording site to the boundary. Recording sites between –0.7 and –0.1 mm (i.e., below the boundary) were assigned to the deep layers, those between 0 and 0.5 mm (above the boundary) were assigned to layer 4 and those between 0.6 and 1.0 mm to the superficial layers. Sites below –0.7 mm and above 1.0mm were excluded from the analysis. Also in this experiment, we excluded recording sites with a VR less than 3. The number of remaining MUA recording sites per compartment were as follows: monkey 4: $N_{deep} = 76$, $N_{layer\ 4} = 97$, $N_{superficial} = 33$; monkey 5: $N_{deep} = 84$, $N_{layer\ 4} = 87$, $N_{superficial} = 31$. Recordings from different penetrations were aligned on the basis of the layer 4c boundary location before averaging across penetrations. To estimate the latency of the CSD modulation, we used the current sink in layer 5, because it was a reliable feature of both the figure enhancement and ground suppression. The current sink was well fit by a Gaussian density function: $g(t) = a \cdot g(t, \mu, \sigma)$ with mean μ and standard deviation σ, and amplitude a. As a measure for the latency of the sink, we took the time point at which the fitted curve reached 33% of its maximum. To quantify the reliability of figure enhancement and ground suppression across the different laminae, we computed d-primes: d_{FU} and d_{UG}, as described earlier. As we were particularly interested in the laminar profile of ground suppression, we only included penetrations in the laminar analyses with significant UGM when averaging across the entire penetration ($P < 0.05$, Wilcoxon signed-rank test). Note that this two-tailed test cannot cause a bias in the results.

Statistical significance of CSD sinks/sources was assessed using a non-parametric bootstrap cluster statistic. The full details are given in Self et al. (2013). Briefly, 2-dimensional (time × depth) t-statistic maps were calculated for each penetration for the difference between figure and uniform, or uniform and ground. These t-maps were thresholded at $P < 0.05$ (two-tailed), and adjacent t-scores above threshold were clustered and the absolute values summed to produce a cluster statistic. Bootstrapping was used to assess the significance of these clusters.

Eye movement analysis

The monkeys had to maintain their eye position within a 1° diameter fixation window. We carried out a stratification analysis to investigate the potential effect of small differences between the eye positions in the figure, ground, and uniform stimulus conditions (Roelfsema et al. 1998; Poort et al. 2012; Self et al. 2013). We computed the average horizontal and vertical eye position in each trial. We then divided the fixation window in 4 x 4 bins of 0.25° × 0.25° and assigned every trial to one of these bins based on the average eye position. We equated the number of trials in each bin across conditions (figure, ground, uniform) by randomly removing surplus trials to ensure that the distribution of eye movements was similar across these conditions and reanalyzed the data of the trials that remained after stratification.

Results

Experiment 1: behavioral task

We trained 3 monkeys to perform a figure-detection task with 2 epochs (see Fig. 1B,C and Materials and Methods). After the monkey directed gaze to the FP, we presented the first stimulus that was either a figure during a period of 400 ms at 1 of 4 possible locations or no figure was presented (uniform condition). This was followed by a second period of 400 ms in which a second stimulus was presented, which could again contain a figure or no figure. At the end of period 2 the FP disappeared. If a figure was present in period 2, the monkey had to make a saccade to its center. If no figure was present, he had to maintain fixation to obtain a reward (catch trial). Note that the stimulus during the first period was uninformative about the required saccadic eye movement although we cannot entirely rule out the possibility of covert eye movement planning during this epoch.

Figure enhancement and ground suppression in V1 and V4

Figure 2A,E shows the activity elicited by the figure, the background and the uniform texture in V1 and V4 during the first stimulus epoch, averaged across 3 monkeys. Before pooling the neuronal responses across the recording sites, we first normalized the activity to the peak response, which is elicited after around 40 ms. This initial response in V1 was similar in the 3 conditions, but after a delay the responses to the figure became enhanced relative to responses to the background and uniform texture (Fig. 2A, blue trace in the lower panel shows the difference between figure and uniform texture, FUM). The modulation of neuronal activity may appear small if it is compared with the initial peak response, but it is in fact quite strong in the later period, when these transients have subsided. The V1 population response elicited by the figure was enhanced by 106% relative to the response evoked by the background (time window 150–300 ms). After an additional delay, the responses to the background became suppressed relative to the uniform texture (UGM; uniform texture minus background response, green in lower panel of Fig. 2A). Note that this later suppression is induced by a figure in the opposite hemifield. Compared with the uniform texture, V1 activity elicited by the figure was enhanced by 42% relative to the response evoked by the uniform texture (Wilcoxon signed-rank test, all monkeys $P < 0.001$), and the response evoked by the ground was reduced by 31% relative to the uniform texture (all monkeys $P < 0.01$).

In V4, the RFs were much larger than in V1 (Motter 2009) and in most cases the V4 RF overlapped with both the interior and the edges of the figure so that the figure can act as a pop-out stimulus at this spatial scale (Roelfsema et al. 2002; Poort et al. 2012). As a result, there was a relatively early enhancement of V4 responses to the figure compared with the responses to the background and uniform texture (Fig. 2E). As in V1, this early enhancement was followed by a delayed suppression of the response to the background relative to the uniform texture, caused by the presence of a figure in the opposite hemifield. When compared with the response elicited by the uniform texture, V4 activity evoked by the figure was enhanced by 30% and the response to the background was reduced by 18% (both $Ps < 10^{-6}$, both monkeys $P < 0.01$). We determined the latency of these effects by fitting curves (see Methods). The visual response latency in V1 was 39 ms with a 95% confidence interval (CI) of 38–40 ms, and it was followed by FUM at 82 ms (CI 68–102 ms)—significantly later ($P < 0.001$; bootstrap analysis)—which was, in turn, followed by UGM at 137 ms (CI 136–141 ms), which was significantly later ($P = 0.02$). The latency of the visual response in V4 was 49 ms (CI 48–50 ms), followed by FUM at 57 ms (53–62 ms), which was in turn followed by UGM at 133 ms (CI 132–140 ms) (latency differences, both $Ps < 0.001$).

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

FUM in V1 was later than FUM in V4 ($P = 0.002$), as shown previously (Poort et al. 2012), but we found that the timing of UGM was similar in areas V1 and V4 (137 ms in V1 vs. 133 ms in V4; $P = 0.88$). Thus, the suppressive effect of the figure in the opposite hemisphere has a similar timing in the 2 areas.

We computed *d*-primes (see Methods) to quantify how reliably individual recording sites discriminated between a figure and a uniform texture (d_{FU}) and between a uniform texture and the background (d_{UG}). Most of the V1 recording sites exhibited an increased response to the figure relative to uniform textures as well as a reduced response to the background (Fig. 2*C*) (d_{FU}, mean 0.16, d_{UG}, mean 0.10, Wilcoxon signed-rank test, both $Ps < 10^{-10}$). In V4 the results were similar because the figure elicited a greater response than the uniform textures, and responses to the background were suppressed relative to those evoked by uniform textures (Fig. 2*G*) ($d_{FU} = 0.90$, $d_{UG} = 0.44$, Wilcoxon signed-rank test, both $Ps < 10^{-6}$).

The discrimination between figure and uniform textures in V4 was stronger than in V1 (V1 $d_{FU} = 0.16$, V4 $d_{FU} = 0.90$, Wilcoxon rank-sum test, $P < 10^{-9}$), and the same was true for the discrimination between uniform textures and background. We computed the correlation between d_{FU} and d_{UG} across recording sites to investigate whether neurons tended to co-express both effects. Interestingly, the correlation between figure enhancement and ground suppression *d*-primes was not significant in V1 and V4 (V1, $r = 0.16$, $P = 0.10$, V4, $r = -0.31$, $P = 0.06$). This result indicates that figure enhancement and ground suppression are separate processes that influence different circuits, as is also evident from the difference in their timing.

In the first stimulus period, we presented a figure–ground display but the monkey was not required to make an eye movement. After 400ms the second stimulus appeared. If a figure was present in the second phase, it served as target for an eye movement. We pooled across all conditions with a particular stimulus in the second period, allowing the stimulus in the first period to vary (Fig. 1*C*). We ensured that the stimulus history was balanced so that the first stimulus did not predict the second stimulus, which enabled us to examine the relative timing of figure enhancement and ground suppression in the 2 stimulus period and determine the possible effect of eye movement planning. When we corrected for the onset time of the second stimulus (at 400 ms), we found that the latency of figure enhancement in V1 (Fig. 2*B*) was 76 ms (CI 66–91 ms) and that it was followed by ground suppression at 141 ms (135–146 ms), significantly later ($P < 0.001$). In V4 (Fig. 2*F*), the latency of figure enhancement was 76 ms (CI 71–83 ms), which was followed by ground suppression at 137 ms (133–146 ms) ($P < 0.001$). Interestingly, the figure enhancement in the second period occurred at similar times in V1 and V4, whereas figure enhancement in V4 preceded figure enhancement in V1 in the first epoch. The main difference in the timing between epochs was a 20 ms delay in the V4 figure enhancement for the second stimulus (57 vs. 76 ms). This extra V4 delay when the perceptual interpretation needs to change (see Fig. 1*C*, the stimulus could change to a figure, ground, or uniform condition) is in accordance with the longer time constants associated with activity changes in higher cortical areas (Chaudhuri et al. 2015). We note, however, some caution is warranted with this interpretation, because the transitions in the RF stimulus also differed between the 2 epochs (gray screen to texture for stimulus 1 and one texture to another texture for stimulus 2, see Methods).

Figure enhancement and ground suppression were also highly consistent across the population of recording sites in the second stimulus period. In V1, the average response elicited by the figure was enhanced by 40% relative to the response evoked by the uniform texture (Fig. 2*B*, all monkeys $P < 10^{-4}$), and the response evoked by the ground was reduced by 31% (all monkeys $P < 0.01$). In V4, figure enhancement was 43%, on average, and ground suppression 16% (Fig. 2*F*, all $Ps < 0.001$). The same result held up when we examined the *d*-primes. Our measure for figure enhancement, d_{FU}, had a mean value of 1.08 in V4, higher than the value of 0.15 in V1 ($P < 10^{-11}$). Similarly, ground suppression in V4 with a mean d_{UG} of 0.31 was stronger than that in V1 with a mean of 0.09 ($P < 10^{-10}$, Fig. 2*D*,*H*). As in period 1, the correlation between figure enhancement and ground suppression *d*-primes was not significant in V1 (V1, $r = 0.16$, $P = 0.11$) and there was even a significant negative correlation in V4 ($r = -0.53$, $P < 0.01$). However, this correlation failed to reach significance when the data of 2 monkeys were analyzed separately (both $Ps > 0.16$). We conclude that neuronal activity in period 2 was remarkably similar to that in period 1, and that the findings therefore do not depend strongly on eye movement planning. In both periods, figure enhancement in V1 and V4 occurred before ground suppression. The strength of figure enhancement was a poor predictor for the strength of ground suppression across neurons, which confirms that figure enhancement and ground suppression are different processes.

Eye movements do not account for figure enhancement or ground suppression

Small differences between the average eye position in the figure, uniform, and background stimulus conditions (within the 1° fixation window) could in principle contribute to the response differences that we observed. We therefore carried out a stratification control analysis in which we first made the distribution of eye position the same across stimulus conditions (see Methods) and repeated our analysis. We found that the neural *d*-prime values after stratification (period 1, V1 $d_{FUstrat} = 0.17$, $d_{UGstrat} = 0.11$, V4 $d_{FUstrat} = 0.94$ $d_{UGstrat} = 0.43$; period 2, V1 $d_{FUstrat} = 0.16$, $d_{UGstrat} = 0.09$, V4 $d_{FUstrat} = 1.03$ $d_{UGstrat} = 0.33$) were similar to the original *d*-prime values without stratification (all Wilcoxon signed-rank test comparing neural *d*-primes before and after stratification $Ps > 0.09$). Thus, small differences in eye position between the conditions cannot account for figure enhancement or ground suppression.

The profile of figure enhancement and ground suppression across the cortical layers

Next, we studied the strength of figure enhancement and ground suppression across the cortical layers of V1 using laminar electrodes in 2 different monkeys (monkey 4 and 5). We presented textures containing a figure to create the figure and background conditions and also uniform textures (Fig. 1*B*). As in Experiment 1, the animals performed a figure-detection task, but now there was only a single epoch. The monkeys either made an eye movement to the figure (on figure/ground trials) or maintained fixation if there was no figure (uniform trials). Relative to uniform textures, figure responses were enhanced (Wilcoxon signed-rank test, both monkeys, $P < 0.001$) and background responses were suppressed (both monkeys $P < 0.001$) (Fig. 3*A*). The magnitude and latency of the ground suppression were similar to that in V1 of the monkeys that participated in Experiment 1. Averaged across the layers, the latency of figure enhancement was 84 ms (CI 68–95 ms) and the latency of ground suppression was 171 ms (127–202 ms). In a previous

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

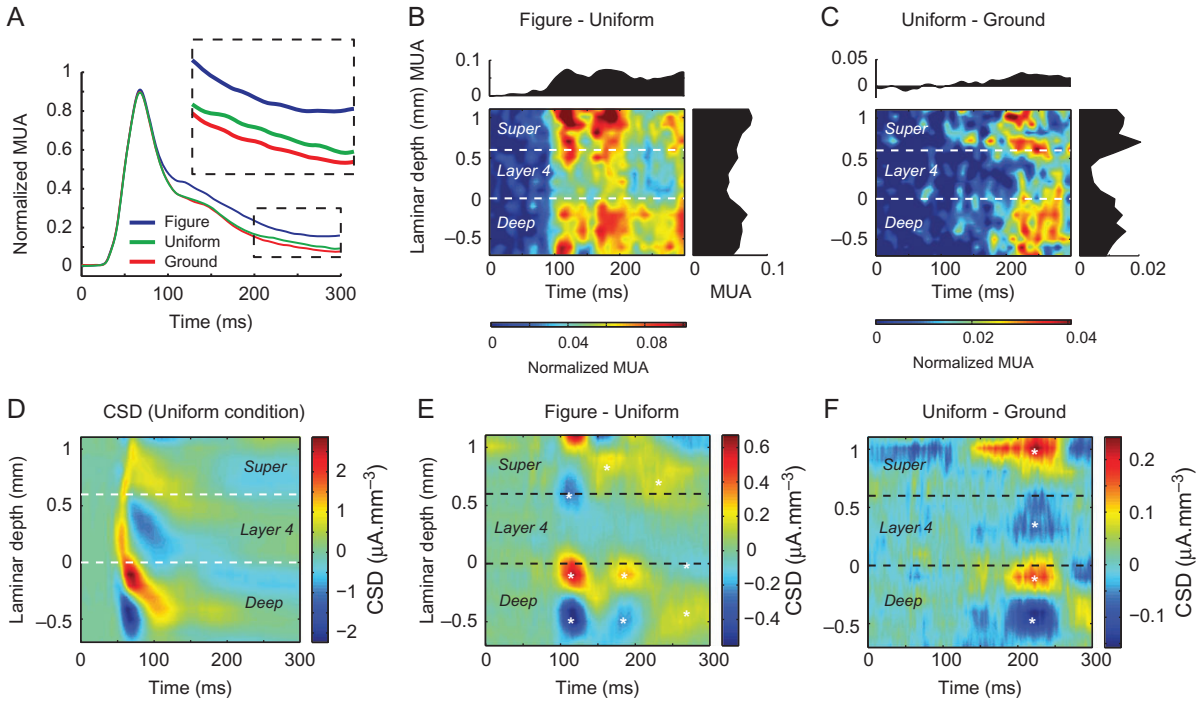

Figure 3. The laminar profile of figure enhancement and ground suppression. (*A*) Average neuronal activity in V1 when the RF was on the figure (blue), background (red) or on the uniform texture (green) (408 recording sites). (*B*) Laminar profile of the difference between MUA evoked by the figure and uniform texture. The boundary between layers 4 and 5 is at 0 mm. (*C*) Laminar profile of the difference between activity evoked by the uniform texture and background. (*D*) CSD at different cortical depths elicited by the appearance of a uniform texture. (*E*) Difference in the CSD evoked by the figure center and a uniform texture. Warm colors show stronger sinks in the figure condition (and/or stronger sources in the uniform condition). The white asterisks mark significant sinks and sources as assessed by a bootstrap cluster statistic (Methods). (*F*) Difference in normalized CSD evoked by the uniform texture and the background. Warm colors show stronger sinks in the uniform condition (and/or stronger sources in the ground condition).

study, we found that FGM had the strongest influence on neuronal activity in the deep and superficial layers, and the weakest influence on activity in input layer 4 (Self et al. 2013). This previous study compared the ground condition to the figure condition, and it did therefore not separate the contributions of figure enhancement and ground suppression.

To isolate figure enhancement, we here compared the responses elicited by the figure to those elicited by the uniform texture (Fig. 3*B*). Figure enhancement was considerably stronger in the superficial and deep layers than in layer 4. For the quantification of figure enhancement, we grouped recording sites into 3 laminar compartments (deep, layer 4 and superficial) and calculated d_{FU} (figure vs. uniform). The level of d_{FU} varied significantly across these laminar compartments (Friedman test, $P = 0.007$). Post hoc tests revealed that the difference between the deep layers and layer 4 was significant (Wilcoxon signed-rank test, $P = 0.02$, with Bonferroni correction for multiple comparisons), and the difference between the superficial layers and layer 4 was significant too ($P < 0.03$). There was no significant difference in figure enhancement between the superficial and deep layers ($P = 0.69$). We then examined the laminar profile of ground suppression by comparing the uniform and ground conditions (Fig. 3*C*). The laminar profile of ground suppression was similar to that of figure enhancement. The values of d_{UG} differed significantly between laminar compartments (Friedman test, $P < 0.001$). Ground suppression was significantly stronger in the deep and superficial layers than in layer 4, and suppression was also slightly stronger in the superficial layers than in the deep layers (deep vs. layer 4: $P = 0.03$, superficial vs. layer 4: $P = 0.004$; deep vs. superficial: $P = 0.04$). We next examined the correlation between figure enhancement (d_{FU}) and ground suppression (d_{UG}) across recording sites, but it was not significant ($r = 0.02$, $P = 0.47$).

To investigate the synaptic contributions underlying these changes in spiking activity, we studied the laminar CSD profile. Sinks in the CSD represent the laminar locations where currents flow into the neurons, and they therefore represent putative excitatory inputs, whereas sources represent the laminar locations where the currents flow out of the neurons (Mitzdorf 1985). The appearance of a full-screen uniform texture produced a typical laminar pattern of current flow with current sinks beginning in layer 4 and then spreading into the superficial and deep layers (Fig. 3*D*). The earliest sinks in layer 4 are thought to represent excitatory feedforward input from the LGN (Self et al. 2013). We next examined the differences in current flow between the figure and uniform conditions, which provides insight into the connections that contribute to figure enhancement (Fig. 3*E*). If the figure fell in the neurons' RF, we observed an extra sink in the upper layers (most likely in layers 1 and 2) and layer 5, at a latency of 97 ms (CI 76–103 ms). This pattern resembles the difference in current flow when we compared the figure condition with the background (Self et al. 2013). Interestingly, layers 1, 2, and 5 are targeted by feedback connections from higher visual areas, which suggests that figure enhancement is caused by excitatory feedback from higher visual areas. To examine the currents underlying ground suppression, we subtracted the CSD when the RFs fell on the ground from the CSD elicited by a homogeneous texture, because the homogeneous texture elicited the strongest MUA response. The laminar profile of this CSD difference was very similar to that

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

underlying figure enhancement, with stronger current sinks in the upper layers and layer 5. Thus, the sinks in the ground condition in layers 1, 2, and 5 were weaker than those elicited by a homogeneous texture, which suggests that ground suppression is associated with a decreased synaptic drive into these layers. The influence of ground suppression on the CSD occurred at a latency of 181 ms (CI 144–194 ms) after stimulus onset, at approximately the same time as the suppression of spiking activity caused by the presence of a figure far from the RF of the neurons.

Discussion

Perceptual organization enhances the representation of figures relative to the background (Driver and Baylis 1996; Baylis and Cale 2001; Peterson and Skow 2008; Salvagio et al. 2012). Researchers call the enhanced representation of figures over the background FGM (Lamme 1995). Here, we studied the neuronal correlates of perceptual organization with electrophysiology in V1 and V4, using a homogenous texture as the neutral condition. We found, for the first time, that both figure enhancement and ground suppression contribute to FGM in both cortical areas. Figure enhancement occurred first in V4 and in V1, and after an additional delay the representation of the background was suppressed in both areas. The difference in the timing between figure enhancement and ground suppression implies that these mechanisms are at least partially independent, and our finding that enhancement and suppression were largely uncorrelated across recording sites in V1 and V4 supported this notion of independence.

Yet, figure enhancement and ground suppression were not dissimilar in all respects. We found that these processes had similar profiles across the cortical layers, with the strongest effects on spiking activity in the superficial and deep layers and the weakest effects in layer 4. Furthermore, figures led to increased sinks in layers 1, 2, and 5 and a stronger source in layer 6 than the uniform texture, and similarly, uniform textures lead to increased sinks/sources in these same layers when compared with backgrounds. Layers 1, 2, and 5 are the targets of feedback connections from higher visual areas, in particular V2 (Rockland and Pandya 1979; Rockland and Virga 1989; Anderson and Martin 2009). This result, therefore, suggests that feedback projections are most active in the figure condition, less active in the uniform condition, and least active in the background condition. We note, however, that these laminar profiles are the result of subtracting the CSD in one condition from that in another condition. Thus, these data are also consistent with the alternative hypothesis that the background causes strong sources in layers 1, 2, and 5, combined with a sink in layer 6. Yet, we do favor the first hypothesis because the background suppression requires the integration of information across large regions of the visual scene (compared with connection schemes relying on only intra-areal lateral connections) (Angelucci and Bullier 2003; Bair et al. 2003). Neurons in higher visual areas, such as V4 and area TEO (Markov et al. 2011), seem a likely source for these feedback effects that are strongest if they respond to figures, because they have large RFs and they send feedback to layers 1, 2, and 5 in lower areas where we found stronger sinks in responses to figures.

The role of response enhancement and suppression in perceptual organization

A number of previous studies investigated the influence of perceptual organization on neuronal activity in the visual cortex. Previous fMRI studies reported that the representations of figures are enhanced (Scholte et al. 2008) that the representation of the background is suppressed (Likova and Tyler 2008) or a combination of both effects (Strother et al. 2012). Important questions were left open by these fMRI studies because they could not separate the representation of figures and background in the higher visual areas, and fMRI may not distinguish between variations in firing rate of the neurons and changes in synaptic input (Logothetis et al. 2001; Viswanathan and Freeman 2007).

Of particular relevance is a previous study that studied spiking activity in V1 and V4 of monkeys during perceptual organization (Chen et al. 2014). The monkeys had to identify a target string of collinear line elements among irrelevant background elements. The string evoked enhanced V1 activity with a latency of around 95 ms, and the activity elicited by background elements was suppressed approximately 20 ms later. This study also demonstrated that V4 responses elicited by the string were enhanced after 59 ms, but again, the V4 responses to the background elements were not measured separately. Gilad et al. (2013) used a similar task design and monitored neuronal activity in V1 with voltage-sensitive dye imaging. Also in this study, neuronal activity elicited by the string was enhanced and activity elicited by the background elements was suppressed, but the authors did not report a significant difference in latency between enhancement and suppression.

Different processing phases during texture segregation

To enhance our understanding of the processes responsible for perceptual organization, we here capitalized on the texture segregation task (Fig. 4). We obtained evidence for a rule of thumb where the latency of an effect on the activity of a V1 cell depends on the relevant spatial scale (Tsotsos et al. 2008). The neuron's first spikes code the features in its RF, including local line orientation (phase 1 in Fig. 4). Early contextual effects near the boundaries between figure and ground follow, and they cause a local enhancement of activity (phase 2). The next phase is the enhancement of the representation of the figure center, involving the integration of features across a few degrees of visual angle (phase 3). In the last phase, figures that are many degrees away from the RF and that can even be in the opposite hemifield suppress neuronal activity (phase 4).

The phases of boundary detection, region filling, and late suppression (phases 2–4 in Fig. 4) require different computations and thus rely on different neuronal mechanisms. Figure boundaries can be detected by local inhibition between neurons with nearby RFs tuned to the same orientation (Grossberg and Mingolla 1985; Li 1999; Itti and Koch 2001). This suppression is present at an early phase of the response (Knierim and Van Essen 1992; Kastner et al. 1997; Levitt and Lund 1997; Bair et al. 2003) and is strong in image regions with a homogeneous orientation and weaker at figure boundaries. It can, therefore, explain the early response enhancement at figure boundaries in V1 (Lamme et al. 1999) and V4 (Poort et al. 2012) as the relative lack of suppressive influences from neurons tuned to the same orientation. Higher areas represent the figure and its boundaries at a coarser resolution (Fig. 4). Pop-out can occur in these areas when the neurons' RF covers the figure so that neurons in the surround tuned to the same orientation are not well driven and provide only little inhibition. It seems likely that this early enhancement of the representation of boundaries is related to 'border-ownership' signals in V1, V2,

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

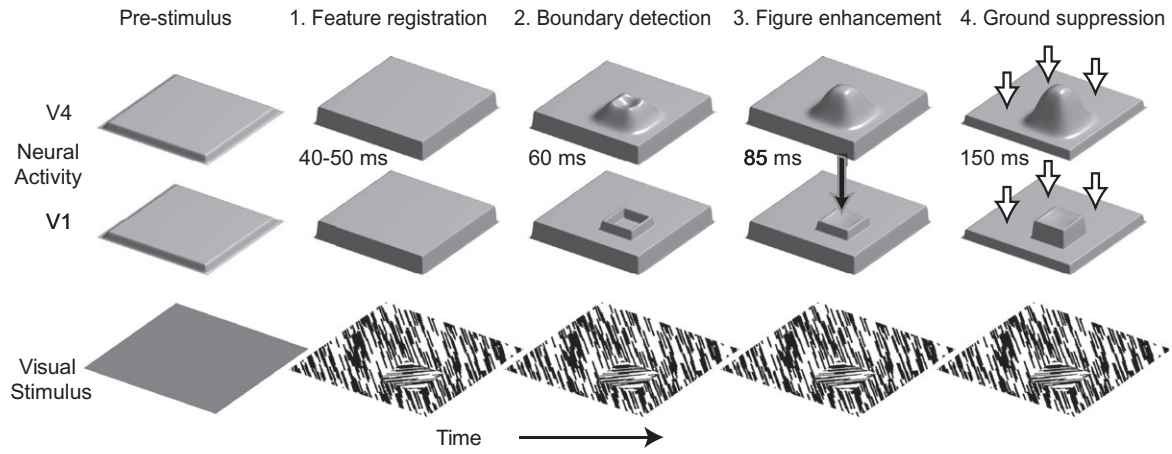

Figure 4. Time course of figure–ground segregation in V1 and V4. First, image features (the local orientation of line elements) are registered. In the second phase, boundaries are detected through a local inhibitory interaction. V4, with its large RFs, represents the figure with a lower spatial resolution than V1 so that edges are more diffused, and the representation of the center of the figure is enhanced at an early point in time (unlike in V1). Third, responses elicited by the center of the figure are now also enhanced in V1. The laminar V1 profile is consistent with a feedback influence from higher visual areas (black arrow). Fourth, the presence of a figure elsewhere appears to reduce feedback from higher visual areas, like area TEO, into layers 1, 2, and 5, resulting in a relatively global suppression of the background representation in V1 and V4 at similar time points (white arrows).

and V4, which code the side of edges that belong to the figure (Zhou et al. 2000; Craft et al. 2007).

The next phase is region filling (figure enhancement in Fig. 4). Now also image elements that are in the center of the figure are labeled with enhanced neuronal activity (Lamme 1995). It is likely that this phase relies on an *excitatory* top-down effect, from neurons in higher areas that represent the figure with extra activity to neurons in lower areas tuned to the same orientation (Poort et al. 2012). Indeed, lesions in higher visual areas reduce modulation at the center but leave boundary modulation intact (Lamme et al. 1998; see also Hupé et al. 1998 who reported that inactivation of area MT had the strongest effect on stimuli of low salience). The present results confirm that region filling increases neuronal activity over the level elicited by homogenous textures. This feedback scenario is also in accordance with the earlier emergence of figure enhancement in V4 than in V1 during texture segregation (Poort et al. 2012) and contour detection (Chen et al. 2014). In this study, FGM in the center of the figure occurred in V4 before V1 in the first stimulus period, but in the second epoch the timing in V1 and V4 was similar. The main difference in timing between the epochs was an increase in delay of V4 FGM in the second epoch, which may be related to a form of inertia of activity of higher visual areas when perceptual representations need to be updated. Indeed, the time constants of neuronal activity in higher areas are longer than those in lower areas (Chaudhuri et al. 2015), although it should be noted that our experiment did not rule out alternative explanations that are related to differences in RF stimulation between the 2 epochs (see Methods).

After yet an additional delay of about 50ms, neuronal activity elicited by the background in V1 and V4 neurons is suppressed by figures that are far from the RF. This 50 ms delay is longer than the 20 ms delay observed in V1 by Chen et al. (2014), which is in accordance with the rule of thumb mentioned above, because the neurons' RFs were farther from the figure in the background condition of this study than in Chen et al. (2014). Furthermore, we here report that the pattern of enhancement followed by suppression also occurs in V4, and that the suppression in V1 and V4 occurs at similar time points. This initial focal response enhancement (Fig. 4, phase 3) followed by delayed global inhibition (Fig. 4, phase 4) could be a general principle that appears to hold true across visual tasks and visual cortical areas. On the one hand, we measured enhancement and suppression for the same recording sites and observed that the strengths of these 2 effects are independent. On the other hand, the laminar profile of MUA and the CSD was similar for enhancement and suppression and suggested that both effects represent influences of feedback from higher visual areas. We mentioned above that we favor the interpretation that excitatory feedback from higher areas is highest for figural image elements, weaker for elements of a homogeneous texture, and weakest for the background. In this view, figure detection would boost representations in higher visual areas and cause extra excitatory feedback at the figure location in early visual cortex, while reducing excitation at other locations, thereby causing ground suppression.

In the section above, we indicated how different processing phases appear at distinct time points of the visual response. The hypothesis that these phases are distinct is inspired by the timing of the response modulations, the effects of attention and lesions, the activity profiles across the cortical layers, as well as by computational considerations. We note, however, that the visual cortical hierarchy is complex and consists of multiple parallel streams with connections that skip hierarchical levels (Felleman and Essen 1991; Markov et al. 2013), and that the entire causal chain of events remains to be fully understood. For example, V4 could directly contribute to figure enhancement in V1 or indirectly through V2. Furthermore, we do not yet know which higher level areas contribute to ground suppression, and information about the impact of horizontal connections on ground suppression is lacking. Future work could address these questions with new methods that enable researchers to monitor (Glickfeld et al. 2013) and manipulate (Inoue et al. 2015) specific neural projections in the circuit that includes V1, V2, V4, and higher areas.

Attention and FGM

Some of the processes for texture segregation are related to selection by visual attention although we did not explicitly test

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

the distribution of attention in this study. In a previous study, we demonstrated that the early phase of texture segregation that gives rise to pop-out and boundary detection is largely stimulus-driven, but that the later region filling process that labels the center of the figure with enhanced activity is reduced if the animal directs attention elsewhere (Poort et al. 2012; see also Roelfsema et al. 2007). This labeling process appears to correspond to object-based attention that is directed to all image elements of the figure (Ben Shahar et al. 2007). The effect of attention reported by Poort et al. (2012) occurred after about 159 ms in V4 and later in V1 (after 204 ms), whereas we did not observe such a timing difference for the suppression in V1 and V4 in this study. Nevertheless, our results are compatible with the hypothesis that the suppression caused by a figure far from the RF is related to a shift of attention away from the ground region and towards the figure. Such a sequence of events would be in agreement with studies showing that shifts of attention start with increased activity for the newly attended item followed by a decrease in activity for nonattended items in monkey visual cortex (Khayat et al. 2006; Busse et al. 2008) and with studies in human visual cortex demonstrating late suppression of activity elicited by nonattended items that are near to a target item in tasks that require spatial scrutiny (Boehler et al. 2009). It is therefore of interest that FGM in Experiment 1 also occurred in the first period when the monkeys could ignore the stimulus. We note, however, that we cannot exclude the possibility that the animals directed attention to the figure, because a similar figure had to be selected for an eye movement at a later point in time.

Even if the ground suppression observed by us is independent of top-down attention, the computational mechanisms that underlie the two processes might be related. The late suppressive effects of attention can be understood in the framework of the selective tuning model (Tsotsos et al. 1995), in which visual input propagates to the top of a hierarchical network where a winner-takes-all selection mechanism inhibits activity at lower levels that is unrelated to the winning stimulus through feedback connections. Such a process would explain a suppressive surround around the attended stimulus. Future studies could examine the joint influence of attention and texture segregation in visual cortex to guidee modeling studies, which could aim to integrate figure–ground segregation and attentional selection into a unified framework.

Conclusion and Outlook

These results combined with the previous work demonstrate that texture segregation relies on a number of different processes that unfold at characteristic time scales. An important goal for future research will be to delineate these distinct processes at the columnar and cellular level, and to identify the inter-areal projections that connect these local circuits. Work in mouse visual cortex has begun to provide insight into how different cell types—in particular interneurons—provide a specific contribution to some of these processes. For example, surround suppression is mediated by somatostatin-positive (SOM) interneurons (Adesnik et al. 2012), and feedback connections can excite SOM cells to increase this suppression and vasoactive intestinal peptide-positive (VIP) interneurons, which inhibit SOM cells, to cause disinhibition (Zhang et al. 2014). It is therefore tempting to speculate that boundary detection in the present texture segregation task depends on SOM cells, with a later top-down input to VIP neurons causing disinhibition for region filling and an even later top-down input to the SOM cells for ground suppression. These separate contributions of different interneuron circuits might also account for the independence of the strength of enhancement and suppression across neurons. Unfortunately, the specific contributions of the different interneuron types in the primate system are less well understood. We anticipate that important progress in this domain can be made with the design of new behavioral paradigms for mice and with the development of transgenic monkeys where the role of the specific cell types and projections can be tested during visual perception (Mitchell et al. 2014).

Funding

NWO (Brain and Cognition grant 433-09-208 and ALW grant 823-02-010 to P.R.R.); European Union (Marie Curie action "ABC," PITN-GA-2011-290011, ERC Grant Agreement 339490 "Cortic_al_gorithms," and 604102, Human Brain Project to P.R.R.). The People Programme (Marie Curie Actions) of the EU's Seventh Framework Programme (Project 332141 to J.P.).

Notes

We thank Kor Brandsma and Anneke Ditewig for biotechnical assistance. *Conflict of Interest*: None declared.

References

Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M. 2012. A neural circuit for spatial summation in visual cortex. Nature. 490:226–231.

Allen HA, Humphreys GW, Colin J, Neumann H. 2009. Ventral extra-striate cortical areas are required for human visual texture segmentation. J Vis. 9:2–14.

Anderson JC, Martin KAC. 2009. The synaptic connections between cortical areas V1 and V2 in macaque monkey. J Neurosci. 29:11283–11293.

Angelucci A, Bullier J. 2003. Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons?. J Physiol Paris. 97:141–154.

Bair W, Cavanaugh JR, Movshon JA. 2003. Time course and time-distance relationships for surround suppression in macaque V1 neurons. J Neurosci. 23:7690–7701.

Baylis GC, Cale EM. 2001. The figure has a shape, but the ground does not: evidence from a priming paradigm. J Exp Psychol Hum Percept Perform. 27:633–643.

Ben Shahar O, Scholl BJ, Zucker SW. 2007. Attention, segregation, and textons: bridging the gap between object-based attention and texton-based segregation. Vis Res. 47:845–860.

Boehler CN, Tsotsos JK, Schoenfeld MA, Heinze H-J, Hopf J-M. 2009. The center-surround profile of the focus of attention arises from recurrent processing in visual cortex. Cereb Cortex. 19:982–991.

Busse L, Katzner S, Treue S. 2008. Temporal dynamics of neuronal modulation during exogenous and endogenous shifts of visual attention in macaque area MT. Proc Natl Acad Sci U S A. 105:16380–16385.

Chaudhuri R, Knoblauch K, Gariel M-A, Kennedy H, Wang X-J. 2015. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. Neuron. 88:419–431.

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.

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

Cohen MR, Maunsell JH. 2009. Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci. 12:1594–1600.
Craft E, Schütze H, Niebur E, von der Heydt R. 2007. A neural model of figure-ground organization. J Neurophysiol. 97:4310–4326.
DeSchepper B, Treisman A. 1996. Visual memory for novel shapes: Implicit coding without attention. J Exp Psychol Learn Mem Cogn. 22:27–47.
Driver J, Baylis GC. 1996. Edge-assignment and figure-ground segmentation in short-term visual matching. Cogn Psychol. 31:248–306.
Felleman DJ, Essen DCV. 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex. 1:1–47.
Gilad A, Meirovithz E, Slovin H. 2013. Population responses to contour integration: early encoding of discrete elements and late perceptual grouping. Neuron. 78:389–402.
Glickfeld LL, Andermann ML, Bonin V, Reid RC. 2013. Cortico-cortical projections in mouse visual cortex are functionally target specific. Nat Neurosci. 16:219–226.
Grossberg S, Mingolla E. 1985. Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. Psychol Rev. 92:173–211.
Hupé JM, James AC, Payne BR, Lomber SG, Girard P, Bullier J. 1998. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature. 394:784–787.
Inoue K, Takada M, Matsumoto M. 2015. Neuronal and behavioural modulations by pathway-selective optogenetic stimulation of the primate oculomotor system. Nat Commun. 6:8378.
Itti L, Koch C. 2001. Computational modelling of visual attention. Nat Rev Neurosci. 2:194–203.
Kastner S, Nothdurft HC, Pigarev IN. 1997. Neuronal correlates of pop-out in cat striate cortex. Vis Res. 37:371–376.
Kato H, Bishop PO, Orban GA. 1978. Hypercomplex and simple/complex cell classifications in cat striate cortex. J Neurophysiol. 41:1071–1095.
Khayat PS, Spekreijse H, Roelfsema PR. 2006. Attention lights up new object representations before the old ones fade away. J Neurosci. 26:138–142.
Knierim JJ, Van Essen DC. 1992. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol. 67:961–980.
Lamme VA. 1995. The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci. 15:1605–1615.
Lamme VA, Rodriguez-Rodriguez V, Spekreijse H. 1999. Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. Cereb Cortex. 9:406–413.
Lamme VA, Supèr H, Spekreijse H. 1998. Feedforward, horizontal, and feedback processing in the visual cortex. Curr Opin Neurobiol. 8:529–535.
Landman R, Spekreijse H, Lamme VA. 2003. Set size effects in the macaque striate cortex. J Cogn Neurosci. 15:873–882.
Legatt AD, Arezzo J, Vaughan HG. 1980. Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: effects of volume-conducted potentials. J Neurosci Methods. 2:203–217.
Levitt JB, Lund JS. 1997. Contrast dependence of contextual effects in primate visual cortex. Nature. 387:73–76.
Li Z. 1999. Contextual influences in V1 as a basis for pop out and asymmetry in visual search. Proc Natl Acad Sci U S A. 96:10530–10535.
Likova LT, Tyler CW. 2008. Occipital network for figure/ground organization. Exp Brain Res. 189:257–267.
Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature. 412:150–157.
Markov NT, Ercsey-Ravasz M, Essen DCV, Knoblauch K, Toroczkai Z, Kennedy H. 2013. Cortical high-density counterstream architectures. Science. 342:1238406.
Markov NT, Misery P, Falchier A, Lamy C, Vezoli J, Quilodran R, Gariel MA, Giroud P, Ercsey-Ravasz M, Pilaz LJ, et al. 2011. Weight consistency specifies regularities of macaque cortical networks. Cereb Cortex. 21:1254–1272.
Merigan WH. 1996. Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. Vis Neurosci. 13:51–60.
Mitchell JF, Reynolds JH, Miller CT. 2014. Active vision in marmosets: a model system for visual neuroscience. J Neurosci. 34:1183–1194.
Mitzdorf U. 1985. Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol Rev. 65:37–100.
Motter BC. 1993. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol.. 70:909–919.
Motter BC. 2009. Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of the V1 surface. J Neurosci. 29:5749–5757.
Peterson MA, Skow E. 2008. Inhibitory competition between shape properties in figure-ground perception. J Exp Psychol Hum Percept Perform. 34:251–267.
Poort J, Raudies F, Wannig A, Lamme VAF, Neumann H, Roelfsema PR. 2012. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. Neuron. 75:143–156.
Rockland KS, Pandya DN. 1979. Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. Brain Res. 179:3–20.
Rockland KS, Virga A. 1989. Terminal arbors of individual "feedback" axons projecting from area V2 to V1 in the macaque monkey: a study using immunohistochemistry of anterogradely transported Phaseolus vulgaris-leucoagglutinin. J Comp Neurol. 285:54–72.
Roelfsema PR, Khayat PS, Spekreijse H. 2003. Subtask sequencing in the primary visual cortex. Proc Natl Acad Sci U S A. 100:5467–5472.
Roelfsema PR, Lamme VA, Spekreijse H. 1998. Object-based attention in the primary visual cortex of the macaque monkey. Nature. 395:376–381.
Roelfsema PR, Lamme VA, Spekreijse H, Bosch H. 2002. Figure-ground segregation in a recurrent network architecture. J Cogn Neurosci. 14:525–537.
Roelfsema PR, Tolboom M, Khayat PS. 2007. Different Processing Phases for Features, Figures, and Selective Attention in the Primary Visual Cortex. Neuron. 56:785–792.
Rubin E. 1915.Visuell wahrgenommene figuren. Copenhagen: Glydendalske.
Salvagio E, Cacciamani L, Peterson MA. 2012. Competition-strength-dependent ground suppression in figure-ground perception. Atten Percept Psychophys. 74:964–978.
Scholte HS, Jolij J, Fahrenfort JJ, Lamme VA. 2008. Feedforward and recurrent processing in scene segmentation: electroencephalography and functional magnetic resonance imaging. J Cogn Neurosci. 20:2097–2109.
Self MW, Kooijmans RN, Supèr H, Lamme VA, Roelfsema PR. 2012. Different glutamate receptors convey feedforward and

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016

recurrent processing in macaque V1. Proc Natl Acad Sci U S A. 109:11031–11036.

Self MW, van Kerkoerle T, Supèr H, Roelfsema PR. 2013. Distinct roles of the cortical layers of area v1 in figure-ground segregation. Curr Biol. 23:2121–2129.

Strother L, Lavell C, Vilis T. 2012. Figure-ground representation and its decay in primary visual cortex. J Cogn Neurosci. 24:905–914.

Supèr H, Roelfsema PR. 2005. Chronic multiunit recordings in behaving animals: advantages and limitations. Prog Brain Res. 147:263–282.

Thompson KG, Hanes DP, Bichot NP, Schall JD. 1996. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. J Neurophysiol. 76:4040–4055.

Tsotsos JK, Culhane SM, Kei Wai WY, Lai Y, Davis N, Nuflo F. 1995. Modeling visual attention via selective tuning. Artif Intell, Special Volume on Computer Vision. 78:507–545.

Tsotsos JK, Rodríguez-Sánchez AJ, Rothenstein AL, Simine E. 2008. The different stages of visual recognition need different attentional binding strategies. Brain Res. 1225:119–132.

Viswanathan A, Freeman RD. 2007. Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. Nat Neurosci. 10:1308–1312.

Zhang S, Xu M, Kamigaki T, Do JPH, Chang W-C, Jenvay S, Miyamichi K, Luo L, Dan Y. 2014. Long-range and local circuits for top-down modulation of visual cortex processing. Science. 345:660–665.

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

Downloaded from http://cercor.oxfordjournals.org/ by guest on September 27, 2016