Behavioral/Cognitive

Resolving the Spatial Profile of Figure Enhancement in Human V1 through Population Receptive Field Modeling

 Sonia Poltoratski[1,2] and Frank Tong[1,2]

[1]Department of Psychology, Vanderbilt University, Nashville, Tennessee 37240, and [2]Vanderbilt Vision Research Center, Vanderbilt University, Nashville, Tennessee 37240

The detection and segmentation of meaningful figures from their background is one of the primary functions of vision. While work in nonhuman primates has implicated early visual mechanisms in this figure–ground modulation, neuroimaging in humans has instead largely ascribed the processing of figures and objects to higher stages of the visual hierarchy. Here, we used high-field fMRI at 7 Tesla to measure BOLD responses to task-irrelevant orientation-defined figures in human early visual cortex (N = 6, four females). We used a novel population receptive field mapping-based approach to resolve the spatial profiles of two constituent mechanisms of figure–ground modulation: a local boundary response, and a further enhancement spanning the full extent of the figure region that is driven by global differences in features. Reconstructing the distinct spatial profiles of these effects reveals that figure enhancement modulates responses in human early visual cortex in a manner consistent with a mechanism of automatic, contextually driven feedback from higher visual areas.

Key words: early visual system; figure–ground; fMRI; LGN; perception; pRF modeling

Significance Statement

A core function of the visual system is to parse complex 2D input into meaningful figures. We do so constantly and seamlessly, both by processing information about visible edges and by analyzing large-scale differences between figure and background. While influential neurophysiology work has characterized an intriguing mechanism that enhances V1 responses to perceptual figures, we have a poor understanding of how the early visual system contributes to figure–ground processing in humans. Here, we use advanced computational analysis methods and high-field human fMRI data to resolve the distinct spatial profiles of local edge and global figure enhancement in the early visual system (V1 and LGN); the latter is distinct and consistent with a mechanism of automatic, stimulus-driven feedback from higher-level visual areas.

Introduction

When we view a scene, our visual system must detect and segment figures from the background environment, guiding our attention toward regions that are likely to contain meaningful objects. Research has shown that local differences in visual feature content (e.g., color, luminance, or orientation) between a figure and its surroundings provide powerful cues for segmentation and detection (Treisman and Gelade, 1980; Bergen and Adelson, 1988; Landy and Bergen, 1991; Nothdurft, 1993).

Conversely, if the figure shares many visual features with its surrounding background, it may be much more difficult to detect; this property of vision is commonly exploited by animals whose coats or skins match their typical environment, making them more difficult for predators to see (Fig. 1A). How might the contextual processing of features at local and larger scales contribute to the visual perception of figures?

Researchers have sought to determine the early visual mechanisms by which figures are differentiated from a surrounding background (i.e., figure–ground perception), with considerable focus on the response properties of neurons in the primary visual cortex (V1) of nonhuman primates (Lamme, 1995; Zipser et al., 1996; Supèr et al., 2001; Marcus and van Essen, 2002; Self et al., 2013, 2019). These studies find evidence of two distinct mechanisms by which figures are differentiated from ground regions in V1: when a feature-defined figure, such as a patch of oriented lines, is presented together with a background region (Fig. 1B), V1 neurons with receptive fields that fall along the boundary between figure and ground exhibit a rapid early enhancement in their response (Fig. 1C, boundary detection). Subsequently, V1 neurons corresponding to the figure region itself exhibit
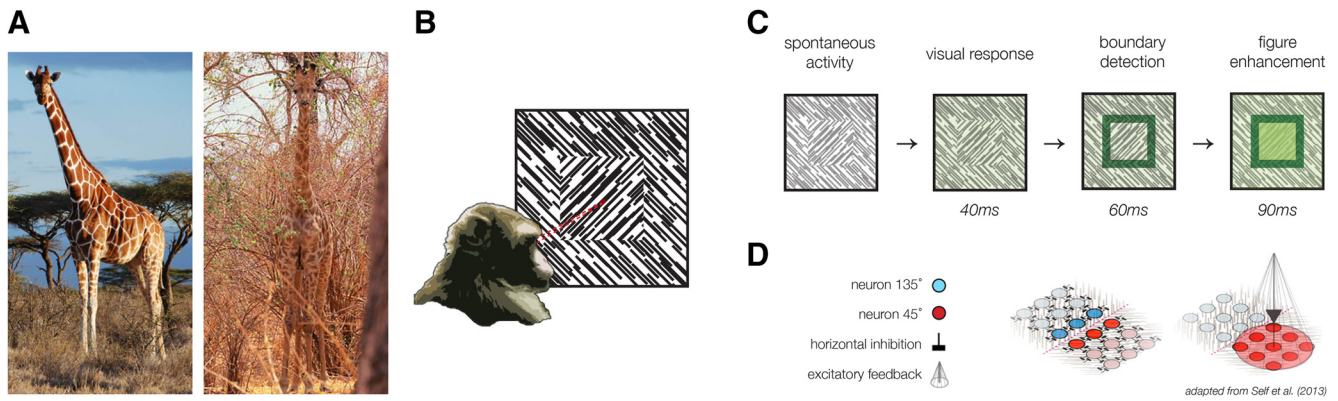
Figure 1. Figure–ground modulation in the wild and in the literature. *A*, An example of animal camouflage, which is pervasive in the natural visual world. Detection of figures is much more difficult when they resemble the color, luminance, and spatial frequency content of the surrounding background. *B*, A sample orientation-defined figure display, typically shown to an animal such that the receptive field of the recorded neuron falls over the figure region. *C*, Proposed mechanisms of figure–ground modulation in the primate V1: after an initial visual response (40 ms), local boundary detection mechanisms enhance responses near the figure–surround boundary (60 ms). Subsequently (90 ms), the entire figure region is enhanced. *D*, The contribution of intrareal horizontal inhibition and excitatory feedback from higher-level visual areas to boundary detection (left) and figure enhancement (right). Adapted from Self et al., (2013) with permission from Elsevier.

response enhancement during the sustained phase of the neural response (Fig. 1C, figure enhancement). The finding of this delayed figure enhancement in primary visual cortex is of particular interest, as it suggests the involvement of more complex neural processes that cannot readily be explained in terms of local feature-tuned surround suppression (Sillito et al., 1995; Bair et al., 2003; Shushruth et al., 2012; Bijanzadeh et al., 2018), particularly by its asymmetrical spatial profile. Indeed, subsequent electrophysiology work has suggested that top–down feedback from higher-order visual areas, including area V4 and the parietal cortex, is likely responsible for inducing figure enhancement responses in V1 (Lamme et al., 1998a; Poort et al., 2012, 2016; Self et al., 2013).

Recently, our group has found that feature-defined perceptual figures evoke enhanced responses throughout the human early visual system, including V1 and the lateral geniculate nucleus (LGN) of the thalamus (Poltoratski et al., 2019). While figure–ground modulation has been well characterized in the early visual cortex of nonhuman primates, in humans the processing of textures, surfaces, and shapes has often been attributed to cortical areas beyond the primary visual cortex, including V4 (Kastner et al., 2000; Kourtzi et al., 2003; Thielscher et al., 2008) and the lateral occipital cortex (Grill-Spector et al., 1999; Vinberg and Grill-Spector, 2008). Our recent work demonstrated the involvement of human primary visual cortex in the enhancement of visual figures (Poltoratski et al., 2019); importantly, we showed that figure–ground modulation in human early visual cortex and LGN is dissociable from effects of directed attention, suggesting an automatic, contextually driven process. However, neurophysiological studies in monkeys have suggested that figure enhancement in V1 becomes severely attenuated if the animal must attend elsewhere to perform a visual task, while border responses remain intact (Schira et al., 2004; Poort et al., 2012). Could it be that our previous observations of enhanced responses to figures in human V1 were caused by local responses specific to the border between the figure and the surround, or can we find evidence of a separate, spatially specific enhancement of the figure region?

One can differentiate between responses driven by the processing of local differences between figure and surround from those caused by a more global figure enhancement by measuring their spatial profiles, which are predicted to be markedly different: boundary responses should appear relatively local to the border between figure and surround, whereas figure enhancement should spread across the entire figure region (Zipser et al., 1996; Self et al., 2019). The prevailing theory of figure enhancement predicts that retinotopic enhancement of activity should be evident throughout the extent of the figure, and moreover, this enhancement should not vary as a function of distance from the boundary (Poort et al., 2012). However, others have challenged the idea that the full extent of the figure may be enhanced in V1, especially at locations far from the boundary (Rossi et al., 2001; Zhaoping, 2003).

Materials and Methods

The goal of our study was to determine whether figure enhancement does indeed occur in human V1 in a manner that can be distinguished from boundary responses, and to determine how the magnitude of this enhancement changes across the figure region as a function of distance from the boundary. We relied on population receptive field (pRF) mapping (Dumoulin and Wandell, 2008; Kok and de Lange, 2014; Wandell and Winawer, 2015) to characterize the response fields of individual voxels to distinguish figure enhancement from local responses elicited by the boundary between figure and surround. Observers performed a visual task (color change detection) at central fixation while task-irrelevant figure–ground stimuli (4° or 6° in diameter) were presented (Fig. 2A, Movie 1). We compared the spatial profile of fMRI responses to the following three types of visual displays: a ground-only condition in which no figure was present; a congruent figure condition in which the figure region is iso-oriented with the background, resulting in a weak figure percept; and an incongruent figure condition in which the figure is orthogonal to the background and appears more perceptually distinct.

While we expected V1 responses to be greater overall for both types of figures than when no figure is presented (ground-only condition), we were particularly interested in how responses to the incongruent figure would differ from the congruent condition in their spatial profile. We recently found that V1 responds more strongly to incongruent figures than to congruent figures, even when attention is directed away from the perceptual figure (Poltoratski et al., 2019; see also Marcus and van Essen, 2002). However, to conclude that this modulation is driven by figure enhancement, it is necessary to resolve the spatial profile of local, spatially constrained boundary responses, which we predicted would occur in response to the congruent figure stimulus, and to demonstrate that an additional enhancement occurs throughout the greater figure region, as we predicted would occur for incongruent figures. Through pRF mapping and the visual reconstruction of fMRI responses to incongruent figures compared with congruent ones, we find compelling evidence that

figure enhancement occurs in human V1 in a manner that can be clearly distinguished from local boundary processing.

Participants. All experiments were performed at the Vanderbilt University Institute for Imaging Science (VUIIS), adhering to the guidelines of the Vanderbilt institutional review board; participants were compensated for their time. Six experienced human observer participants (four females; age range, 23–31 years) were scanned in the main experiment; three of these participants (all females) also completed an additional control experiment, in which larger figures (6° in diameter) were presented. S.P. participated in all experiments.

Scanning procedure. All functional data were collected at the VUIIS research-dedicated 7 Tesla Philips Achieva scanner using a quadrature transmit coil in combination with a 32-channel parallel receive coil array. BOLD activity was measured using single-shot, gradient-echo echoplanar T2*-weighted imaging, at 2 mm isotropic voxel resolution (40 slices; TR, 2000 ms; TE, 35 ms; flip angle, 63°; FOV, 224 × 224; SENSE acceleration factor of 2.9; phase encoding in the anteroposterior direction). Each MRI session lasted 2 h, during which we acquired the following images: (1) one to two functional localizer runs using a central flickering checkerboard to identify retinotopic regions in visual cortex and LGN that corresponded to the stimulus location; (2) seven to eight fMRI runs to measure BOLD activity during the experiment; and (3) five to eight fMRI runs to map population receptive fields in these voxels. Each of the run types lasted 4–6 min.

Scanning procedure for retinotopy. Prior to participating in the functional experiments, each participant underwent a separate session of retinotopic mapping. We used a typical phase-encoded design (Engel et al., 1997; Wandell et al., 2007) in which subjects fixated while they viewed flickering checkerboards consisting of rotating wedges to map polar angle and expanding rings to map eccentricity (Swisher et al., 2012). Retinotopy data were acquired at the VUIIS using a Philips 3T Intera Achieva MRI scanner equipped with an eight-channel receive coil array, using 3 mm isotropic resolution (TR, 2 s; TE, 35 ms; flip angle, 80°; 28 slices; FOV, 192 × 192).

fMRI preprocessing. Data were preprocessed using FSL and Freesurfer tools (documented and freely available for download at http://surfer. nmr.mgh.harvard.edu), beginning with 3D motion correction and linear trend removal, followed by slice-timing correction for pRF runs and a high-pass filter cutoff of 60 s. Functional images were registered to a reconstructed anatomic space for each subject; this registration was first automated in FSL and then checked and corrected by hand. This allowed for the alignment of the fMRI data to the retinotopy data, which were collected in a separate session. The functional localizer data were spatially smoothed using a 1 mm Gaussian kernel to improve the spatial contiguity of delineated regions of interest; no spatial smoothing was performed on the experimental or pRF mapping runs. Further analyses were conducted using a custom Matlab processing stream. The intensities of each voxel were normalized by the mean of the time series, converting to the mean percentage signal change within each run. Outliers were defined as time points for which the response of the voxel measured more than three times its SD from its mean, and were winsorized (Hastings et al., 1947). This condition-blind preprocessing step minimizes the impact of rare spikes in MR intensity while preserving the temporal structure of the responses in each voxel.
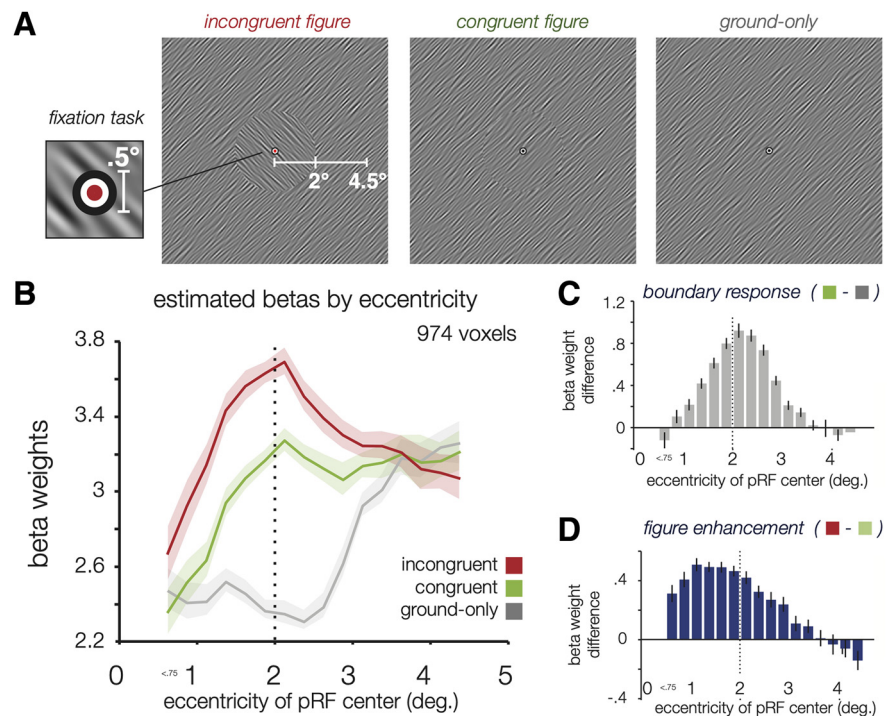

Figure 2. Boundary responses and figure enhancement as a function of eccentricity. *A*, Examples of spatially filtered oriented-noise stimulus displays used in the experiment 2. The following three main conditions are depicted: left, incongruent figure, in which the figure and surround were orthogonally oriented, creating both a perceptual border and an orientation difference; middle, congruent figure, in which the figure and surround are iso-oriented but sampled from distinct patches of noise, creating a visible phase-defined border but a weaker figure percept; right, ground-only condition, in which a single oriented texture fills the visual field. The spatial frequency depicted here is much lower than the experimental stimuli (0.5–8 cycles/°) to improve visibility. See also Movie 1. *B*, V1 responses plotted using mean beta weights for each experimental condition as a function of the eccentricity of the pRF center of each voxel. Dotted line at 2° eccentricity indicates the figure–surround border; error bars depict ±1 SEM across voxels.. *C*, V1 BOLD beta weight difference between congruent figure and ground-only conditions, associated with a predicted boundary response. *D*, V1 BOLD beta weight difference between incongruent and congruent figure conditions, associated with a predicted figure enhancement due to differences in orientation. The spatial profile of this effect as a function of eccentricity is clearly distinct from that of the top panel and appears to extend through the full eccentricity range of the figure but to decline beyond the 2° figure–surround border.
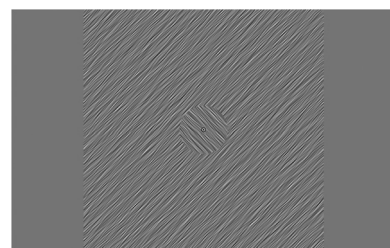
Movie 1. Illustration of the three main experimental stimuli conditions, as follows: incongruent figure, congruent figure, and ground-only condition. Timing is abbreviated for the illustration: in the experiment, stimulus blocks were 16 s long, and were interspersed with 16 s fixation-rest periods. Orientations (45° and 135°) were counterbalanced across blocks, and timing was randomized for every experimental run and participant. Participants reported the brief 200 ms color change at fixation throughout the entire run. [View online]

In all functional experiments in this study, we relied on an fMRI block paradigm and presented blocks of visual stimulation (16 s duration) interleaved among 16 s fixation rest periods. The amplitude of the BOLD response during each stimulus block was then estimated using the general linear model for each voxel. These estimates were averaged across blocks by experimental condition to yield voxelwise mean standardized beta values in each condition.

ROI localization and voxel selection. To initially define retinotopic visual areas V1–V3, each subject participated in a separate retinotopic mapping scan. Boundaries between retinotopic areas V1–V3 were delineated

by hand, by identifying reversals in the phase of the polar angle map measurements. The resulting labels were aligned to the functional space of the current experiment using FSL and Freesurfer software, and this registration was checked and corrected by hand. Additionally, one to two runs of a functional localizer, described below, were collected in the main experimental sessions to identify the LGN in each subject and to select regions of interest (ROIs) responsive to our experimental display from the V1–V3 retinotopic areas. The localizer consisted of blocks of a flickering checkerboard stimulus spanning the full 9° diameter field of view and was designed to yield a large ROI that could be then refined by pRF model fitting.

For all analyses, we used these functional labels in conjunction with the pRF fitting results to define regions of interest. For each subject, all voxels in each visual area were fitted with the pRF model, as described below. For further analyses, we used voxels whose pRF centers were within the range of the mapping stimulus (0.25° to 4.5° eccentricity) and were >0.1°; this limit trimmed instances in which the model predicted nearly no visual response to the mapping stimulus. Following this trimming procedure, we selected the top 33% of best-fitted voxels for each subject in each ROI, as indexed by the R^2 between observed and predicted data. In V1, this yielded fits with R^2 cutoffs that ranged from 0.62 to 0.81 in individual subjects (mean 0.71); corresponding V1 ROIs for each subject ranged from 131 to 187 voxels bilaterally (mean 162). In the LGN, R^2 cutoffs ranged from 0.16 to 0.27 (mean 0.22), yielding ROIs that were 17–36 voxels in size (mean 26.5).

Population receptive field mapping. pRFs correspond to the location in visual space that best drives activity in the population of neurons in each voxel (Dumoulin and Wandell, 2008; Wandell and Winawer, 2015). Unlike standard retinotopic mapping, pRF mapping resolves not only the central location that best drives responses in each voxel, but also the spatial extent of this response field. In each fMRI experimental session, we mapped population receptive fields in retinotopic areas V1–V3, using a 2D circular Gaussian model of pRF structure. pRF modeling involves presenting a mapping stimulus that spans the visual field over time, and estimating the pRF parameters that most likely produced the measured BOLD response in a particular voxel. These parameters define the location, size, and gain of the pRF. pRF properties for each voxel are assumed to reflect the combined RFs of the neural population in a voxel, and appear well aligned with single-neuron receptive field properties, such as contralateral preference, increasing size at greater eccentricities, and increasing size as one ascends the visual hierarchy (Dumoulin and Wandell, 2008; Wandell and Winawer, 2015). Here, we mapped pRFs using a traveling bar stimulus composed of rapidly presented, full-color objects embedded in pink noise (developed by Kendrick Kay, http://kendrickkay.net/analyzePRF/; object stimuli are from Kriegeskorte et al., 2008). The bar stimulus swept through a circular region with a 4.5° radius, the maximal visible field of view at our 7 T scanner, and each mapping run lasted 5 min. Voxelwise responses for each visual area were fitted with a 2D Gaussian pRF model using a custom Matlab pipeline. We used a dual-stage multidimensional nonlinear minimization (Nelder–Mead) fitting procedure: each voxel was initially fitted in a downsampled stimulus space with a fixed Gaussian σ (1°), and then these parameters were used to initialize a full model fitting in native stimulus space. Estimated parameters described pRF position (X, Y), size (σ), and response amplitude of each voxel; we can additionally convert these parameters to measures of polar angle, eccentricity, or full-width at half-maximum (FWHM) to convey pRF size.

Reconstruction of spatial profiles of figure–ground modulation. For the spatial profile visualization, pRF mapping was used to project the differential responses evoked by the figure–ground stimuli to stimulus space. Previous studies (Kok and de Lange, 2014) have adopted an approach of projecting the weighted activity of each voxel into image space by scaling its Gaussian pRF by the response of the voxel to the display, and then calculating the linear sum of all pRFs in a region. In our work, we have found that a multivariate regression-based approach leads to sharper projections. Specifically, we used ridge regression (Hoerl and Kennard, 1970), and, as seen in Figure 7, we model the voxelwise responses in each region as a linear product of pRFs of those voxels and the spatial profile of perceptual effects driving these responses. The spatial profile is

thus solved as a matrix of predictors (β) in the regression, with each weight representing a portion of the original stimulus space (downsampled to 0.25° per weight/pixel, or four pixels per degree; the pRFs are similarly subsampled). The resulting regression can be described as a standard linear model: $y = X\beta + \varepsilon$, where $y \in R^N$ represents the vector of voxel responses, $y \in R^{N \times p}$ consists of a 2D matrix of the mapped Gaussian pRF of each voxel, $\beta \in R^p$ are the predictors (pixels) to be estimated, ε is the error term, and p and N are pixels and voxels, respectively. Since there are many more predictors (pixels) than voxel responses to be predicted, this leads to an inverse mapping problem, for which standard multivariate regression cannot find a unique solution. To constrain the estimation procedure and to minimize overfitting, regularized ridge regression can be used to estimate the predictors for the contrast of each pixel. Whereas linear regression seeks to minimize the sum of squared residuals, ridge regression, or L2 regularization, applies an additional penalty term based on the sum of squared weights (β) to be estimated, thereby giving preference to solutions with smaller β values (Hoerl and Kennard, 1970), as follows:

$$\min_{\beta} \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - x_i^T \beta \right)^2 + \lambda \| \beta \|_2^2,$$

where λ is the scaling factor of the penalty term. Here, we performed ridge regression over the data of each individual ROI and subject to produce projection images. The λ penalty term for each regression was chosen from the range 0–5000 (in increments of 25), using 10-fold cross-validation to find the lowest root-mean-square error (RMSE); in this implementation, a λ of zero corresponds to nonregularized regression. The resulting images were z-scored within subjects, averaged, and smoothed with a small Gaussian kernel ($\sigma = 0.75°$) to create the spatial profile images seen in Figure 7.

Code accessibility. Following the publication of this article, the code for pRF fitting and regression-based visualization will be made publicly available at https://github.com/soniapolt/pRF-figureGround.

Experimental displays. In all experiments, observers viewed orientation-defined figures. To minimize the prevalence of high spatial frequency energy artifacts between the figures and surround, we used oriented bandpass-filtered noise (100% contrast, bandpass filtered from 0.5 to 8 cycles/° with a center orientation of 45° or 135°, ~20° FWHM in the orientation domain). Each orientation stimulus was dynamically generated from random white noise, bandpass filtered in the Fourier domain, convolved with a small Gaussian kernel (in the Fourier domain) to minimize Gibbs ringing artifacts, and then converted back to the image domain. The experimental conditions were presented in 16 s stimulus blocks, during which the oriented noise patterns were dynamically regenerated every 200 ms; these stimulus blocks were interspersed with 16 s fixation–rest blocks. A movie of the experimental conditions for the main and control experiments is available (Movie 1).

Experimental design: main experiment. Sample displays for each of the three conditions in this experiment are illustrated in Figure 2A. In both the incongruent figure and congruent figure conditions, observers viewed a single, centrally presented figure (4° in diameter). For the congruent condition, figure and surround images were generated from two different noise patterns to create a phase-misaligned figure–surround display. In the main experiment, the center and surround directly abutted (but see Control experiment A).

In this experiment, participants passively viewed the figure–surround displays while they performed a color change detection task at fixation throughout the entire experimental run. The fixation changed from black to red for increments of 200 ms at random time intervals, occurring on average of four times per 16 s experimental or fixation block. Participants reported these events by pressing a key on an MR-compatible button box; the average percentage correct across participants was 94.5% (SD, 5.8%). The task difficulty was titrated for this experiment to avoid load-based suppression around fixation, which we have observed in prior studies (Cohen and Tong, 2015) when more difficult central tasks are used. The figures were not task relevant at any point in the main experiment or in the controls; accordingly, we saw no difference in
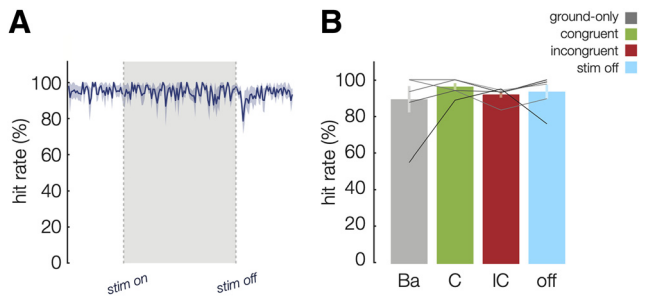
Figure 3. Behavioral performance in the main experiment. ***A***, Behavioral hit rate relative to the onset and offset of 16 s stimulus blocks, plotted in 200 ms increments. Performance is averaged across blocks; shaded error bars indicate ±SEM across subjects. There is no apparent difference between task performance when the stimulus is on the screen versus off, save for a slight dip immediately after the stimulus block that is likely attributable to increased blinking. ***B***, Average hit rate for targets coinciding with the three main experimental condition stimulus blocks, and stimulus-off blocks. Error bars indicate ±SEM across subjects, and grayscale lines correspond to individual participants' performance (some overlap). There is no significant effect of stimulus condition on performance of the fixation task ($F_{(3,20)} = 0.46$, $p = 0.72$), nor is there an overall difference between stimulus-on and stimulus-off performance ($t_{(5)} = 0.59$, $p = 0.58$). While the fixation task in these experiments is not particularly challenging, we would expect some variation in performance if the task-irrelevant figures were actively or preferentially attracting subjects' attention.

task performance relative to the onset/offset of stimulus blocks (Fig. 3), nor were subjects impaired at the task during any of the stimulus conditions relative to blank periods ($F_{(3,20)} = 0.46$, $p = 0.72$). Overall, we see no positive evidence that participants' attention was drawn away from the task by the figures, either generally (Fig. 3A) or in a pattern that would predict the observed enhancement of orientation-defined figures (Fig. 3B).

Control experiment A: gap inserted between figure and ground. In the same session ($N = 6$) as the main experiment, participants also viewed oriented figures that were encircled with a 0.5° grayscale gap (2.0–2.5° eccentricity; see Fig. 8A; Movie 2). This created a visible border around both the incongruent and congruent figures; importantly, the boundaries around these two types of figures were better matched in terms of the local V1 responses they would be expected to evoke compared with those in the main experiment (which were either phase defined or orientation and phase defined). This allowed us to determine whether a difference in local stimulus energy could explain the spatial profile of the effects in the main experiment and to test whether a visible boundary was sufficient to produce figure enhancement when the figure orientation did not differ from the surround. Experimental timing, task, and design were identical to those in the main experiment.

Control experiment B: larger figures. Three subjects from the original experiment performed an additional control experiment in a second session. This control experiment sought to evaluate whether larger figures would still lead to figure enhancement in V1 in regions corresponding to the center of the figure in a manner that could not be readily explained by local effects of feature-tuned surround suppression. This experiment followed the same procedures as in the main study, but used figures that were 6° in diameter. Given our limited field of view in the 7T scanner (~9° in diameter), this stimulus precluded us from measuring many voxels that responded primarily to the surround, but allowed us greater measurement of a range of voxels whose pRFs were within the figure.

Experimental design and statistical analyses. As described in the fMRI preprocessing section above, these experiments relied on the estimation of voxelwise mean standardized beta values in each experimental condition. These estimates are performed within subjects and in native subject neuroanatomical space. For some voxelwise visualizations (e.g., beta estimates as a function of eccentricity; Fig. 2), voxels are pooled across subjects and binned as described in Figure 4. Primary statistical testing is conducted using repeated-measures within-subjects ANOVA, using the mean beta responses in ROIs
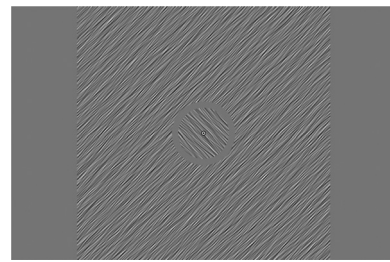
Movie 2. Illustration of the experimental conditions of control experiment A, in which a 0.5° grayscale gap was introduced. As in Movie 1, timing is abbreviated for this illustration. [View online]

defined by the location of the pRF envelopes of the voxels (see Fig. 6). Two-tailed *t* tests were used to further refine our understanding of significant ANOVA results.

Results

Distinct effects of boundary responses and figure enhancement in V1–V3

Figure 2B shows V1 BOLD responses in each of the experimental conditions binned by the eccentricity of the estimated pRF center for each voxel. Bins are 0.25° wide and are thresholded to contain at least 10 voxels pooled across subjects (Fig. 4, details of the binning procedure). fMRI responses were generally stronger in the congruent condition than in the ground-only condition, especially around 2° eccentricity, which corresponded to the location of the boundary between the figure and surround. However, fMRI responses were even greater in the incongruent condition, and this differential response could be observed even in voxels with pRFs located near the center of the figure. It should be noted that the spatial extent of our visual display at the 7 Tesla scanner prevented us from examining the responses of voxels whose pRF centers were located much beyond 4° eccentricity; nevertheless, a modest trend of weaker responses in the incongruent figure condition can be seen, consistent with previous reports that perceptual figures can lead to a suppression of neural responses to the ground region (Appelbaum et al., 2006; Poort et al., 2016; Self et al., 2019).

Not surprisingly, responses to our noise pattern stimuli tended to vary in amplitude by eccentricity, even when no figure was presented (ground-only condition). It is well documented that visual cortical sensitivity to spatial frequency (Singh et al., 2000; Henriksson et al., 2008; Aghajari et al., 2020), temporal frequency (Himmelberg and Wade, 2019), and motion energy (Johnston and Wright, 1985; Kwon et al., 2015) varies systematically as a function of eccentricity in the visual field. Moreover, our behavioral task required participants to focus their attention on the small central fixation point, which may have led to some suppression of neural responses in the neighboring periphery (Tootell et al., 1998; Somers et al., 1999). Because of these factors, the ground-only condition serves as a strong baseline against which to measure responses to the figure stimuli, and comparisons between conditions (Fig. 2C) may be more informative than the absolute magnitudes of the BOLD signal.

We calculated the difference in fMRI response amplitude for congruent figures minus ground only (Fig. 2C, boundary response); this revealed enhanced V1 responses centered at the boundary between the figure and surround, as predicted. A modest degree of skewness toward further eccentricities was also observed, consistent with the fact that pRF sizes generally increase with eccentricity (Fig. 4A). This boundary response, or
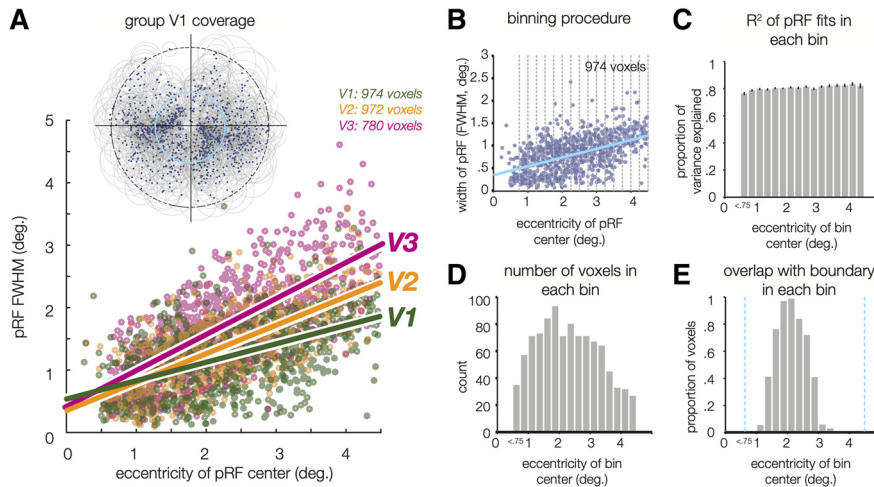
Figure 4. Details of the pRF mapping and binning procedure used in V1 analyses of Figure 2. *A*, Estimated pRF size (as measured by Gaussian FWHM) plotted as a function of eccentricity for all voxels used in the analyses for the main experiment. pRF size increases both as a function of eccentricity and as one ascends the visual hierarchy from V1 to V3. Inset, Total pRF coverage of V1 across subjects; blue dots mark pRF centers, and gray outlines mark FWHM. Blue circles indicate the figure location (2° eccentricity) and the full display extent (4.5° eccentricity). *B*, Scatterplot of voxel pRF size (Gaussian FWHM) as a function of eccentricity, with gray lines indicating bin edges. Bins were 0.25° wide with the exception of the most foveal, which included all voxels with a center pRF eccentricity <0.75°. *C*, Goodness-of-fit of the pRF model across voxels in each bin. Error bars indicate ±SEM across subjects. *D*, Number of voxels in each bin, across subjects. *E*, Proportion of voxels in each bin that overlap with the boundary (as defined in Fig. 5*A*), indicating clear differentiation of voxels in the figure, boundary, and surround regions.

sensitivity to a phase-defined border, appears distinct from the predicted spatial effects of figure enhancement, as it does not spread inward toward the center of the figure.

To determine the spatial profile of the additional enhancement caused by presenting a more distinct figure that differs in orientation from its surround, we calculated the difference in fMRI response for incongruent figures minus congruent figures (called figure enhancement); this revealed enhanced activity that appeared to extend throughout the full 2° radius of the figure (Fig. 2*D*). However, as some variability in pRF size is present at each eccentricity (Fig. 4*A*), it is important to consider whether these effects are primarily driven by voxels that, while centered within the figure, nevertheless receive input from the location of the boundary. To address this question, we performed another analysis that took into consideration both the center position and spatial extent of voxel pRFs. Specifically, we evaluated whether the greater response to incongruent figures might be driven by voxels whose pRFs are centered within the figure but also extend over the boundary, in which case local processing of orientation differences might underlie the apparent effect of figure enhancement. We binned the response of voxels according to whether the most sensitive central region of their pRF, defined by the FWHM envelope, was contained within the figure (negative values of "overlap" with the boundary) or extended past the boundary (positive values of overlap), as illustrated in Figure 5. While the predicted boundary response (congruent figures minus ground only) occurs primarily in voxels whose pRFs overlap with the boundary (Fig. 5, gray curve), the magnitude of the figure enhancement effect (incongruent minus congruent conditions; Fig. 5, blue curve) remains stable and positive for voxels whose response region (FWHM) falls within the figure and does not overlap with the boundary ($F_{(1,3)} = 0.50$, $p = 0.53$).

We observed this same pattern of results in a separate control experiment using larger 6° diameter figures. Participants again performed a visual monitoring task on the central fixation point (0.5° diameter), so both figure and ground stimuli were task irrelevant. The greater response to incongruent than to congruent figures was evident even in voxels whose pRF response fields (FWHM) were >2° away from the figure–surround boundary (Fig. 5*C*; $F_{(1,6)} = 1.66$, $p = 0.25$). In contrast, our measure of the boundary response (congruent minus ground only) was again local to the border, and showed no evidence of spreading toward the center of the figure. Together, these results imply that the enhancement stemming from the incongruent orientation of the figure does not arise from strictly local processing of feature contrast, as such effects would be expected to decline as a function of distance from the boundary. Instead, the heightened response we observe for incongruent figures, compared with congruent figures, is consistent with the predicted effects of figure enhancement in V1.

Finally, we used pRF measurements to identify voxels within V1 whose FWHM envelopes were either fully contained within the figure (Fig. 6*A*, orange), overlapped with the boundary (Fig. 6*A*, magenta), or fell in the surround region beyond the boundary (Fig. 6*A*, purple). Mean BOLD time courses and beta estimates of fMRI response amplitude revealed a distinct pattern of results for these ROIs. The figure ROI showed a significantly greater response to incongruent figures than to congruent figures ($t_{(5)} = 7.86$, $p = 5.4 \times 10^{-4}$), consistent with the predicted effects of figure enhancement, whereas it did not show a boundary effect ($t_{(5)} = 0.88$, $p = 0.42$).

The boundary ROI exhibited a much stronger response to congruent figures than to ground only ($t_{(5)} = 7.7$, $p = 7.4 \times 10^{-4}$), consistent with our prediction that a border response would be evoked by the phase-misaligned iso-oriented figure. The boundary ROI also showed a significantly greater response to incongruent than to congruent figures ($t_{(5)} = 5.46$, $p = 2.8 \times 10^{-3}$), consistent with our expectation that boundary ROI voxels should show some degree of figure enhancement, given that part of their population receptive fields fall on the figure region. A direct comparison indicated that the magnitude of the boundary response was significantly greater than the magnitude of figure enhancement in these boundary–ROI voxels [one-tailed $t_{(5)} = 3.29$, $p = 0.011$; mean beta differences of 0.74 (SE, 0.10) and 0.35 (SE, 0.06), respectively], consistent with findings reported in studies of nonhuman primates (Poort et al., 2012; Self et al., 2019). Finally, the surround ROI did not show any significant differences between the experimental conditions, further demonstrating the spatial specificity of figure–ground modulation. Together, these results provide strong support for the notion that the mechanisms of figure enhancement and sensitivity to the figure–surround boundary lead to distinct patterns of activation in human V1.
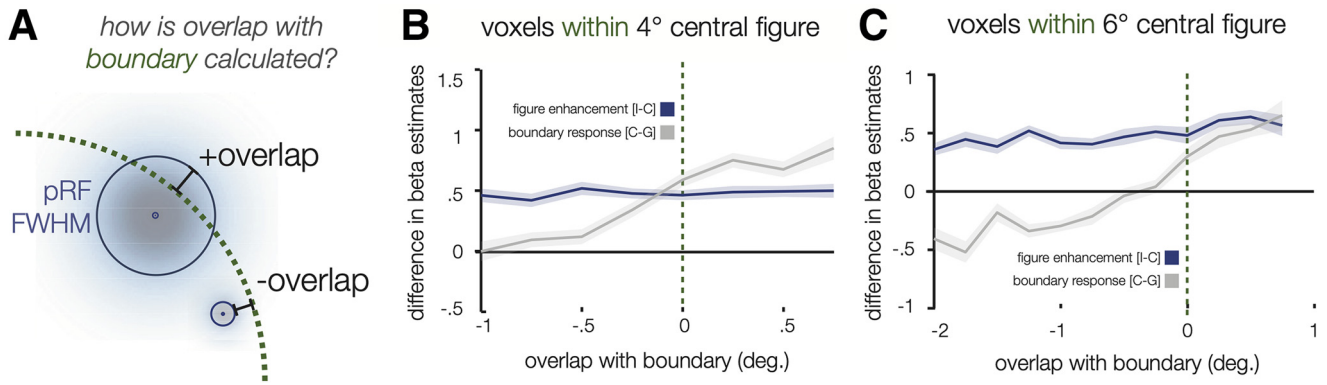
Figure 5. Boundary response and figure enhancement in voxels centered within the figure, plotted as a function of pRF extension beyond boundary. **A**, As illustrated, two pRFs with centers at the same eccentricity within the figure might, by virtue of differences in their size, lead to differential sensitivity to the location of the figure–surround boundary. We calculated whether the central FWHM region of each pRF fell short of the boundary (negative overlap values) or extended past the boundary (positive overlap values). **B**, V1 boundary responses and figure enhancement as a function of distance; while boundary responses (congruent minus ground-only condition, gray) is evident primarily in voxels whose pRFs overlap with the figure/surround boundary (overlap >0), figure enhancement (incongruent minus congruent, blue) occurs along the full extent of measured distances. Bins are 0.25° wide, and those containing <10 data points were trimmed. **C**, Results from a control experiment using 6° diameter figures. While boundary responses again primarily occur in voxels whose pRFs overlap with the boundary (overlap <0), figure enhancement persists even in voxels with up to 2° spatial separation with the figure–surround boundary.

Figure 6. Average BOLD responses across figure, boundary, and surround-selective voxels. **A**, V1 pRF locations of one representative subject (circles, FWHM), illustrating how pRFs were sorted to create ROIs with voxels primarily responsive to the figure (orange, 205 voxels across participants), surround (purple, 293 voxels), or on the boundary (magenta, 476 voxels). Dotted lines denote the spatial extent of our mapping stimulus (9° diameter) and the figure region (4° diameter). **B**, V1 results for these ROIs; left panels depict average BOLD time courses in each condition, with dotted lines marking the onset and offset of 16 s stimulus blocks. Right panels depict averaged estimated GLM beta weights in each condition. Error bars indicate ±1 SEM between subjects, and asterisks indicate significance at $p < .05$, n.s. = no significance. **C**, Results from voxels in the LGN that overlapped with the center, surround, and boundary.

Reconstructing differential responses to figures in stimulus space

pRF modeling also provides an opportunity to visualize the pattern of voxel responses for a given visual area in the native stimulus space (Kok and de Lange, 2014). Here, we used regularized linear regression to infer the spatial profiles of both boundary responses and figure enhancement responses in stimulus space. As described in Figure 7, this approach assumes that a voxelwise vector of response amplitudes can be modeled as a linear product of the pRFs of those voxels and the spatial profile of the differential response in stimulus space. To improve the tractability of the model, we downsampled the effective resolution of the stimulus space to four pixels per degree of visual angle, such that each pixel in the reconstructed spatial profile (Fig. 7) corresponds to 0.25° of the original display.

We used an L2 (ridge) penalty term to minimize overfitting (Hoerl and Kennard, 1970) and a 10-fold cross-validation to select the penalty value that produced a minimal RMSE. This regression method can yield better spatial resolution in the visualization than simply scaling the 2D Gaussian derived from the pRF of each voxel by its response magnitude and averaging the weighted responses of all voxel pRFs throughout the visual field, as the latter approach is sure to introduce some blurring of the reconstruction given the spatial spread of the Gaussian pRFs themselves. We performed this analysis on the differential response to congruent figures versus ground only and to incongruent versus congruent figures, normalized the resulting images within subjects, and then created averaged visualizations across subjects, as shown in Figure 7. The spatial profile of these visualization results clearly support the distinct effects of boundary responses and figure enhancement, and go on to show that the response pattern evoked by incongruent figures, compared with congruent figures, is well described by enhanced representation specific to the figural region.

A clearly visible boundary is not sufficient for figure enhancement in V1

We have demonstrated that orientation-defined figures produce enhanced BOLD responses across the figure region in V1, and
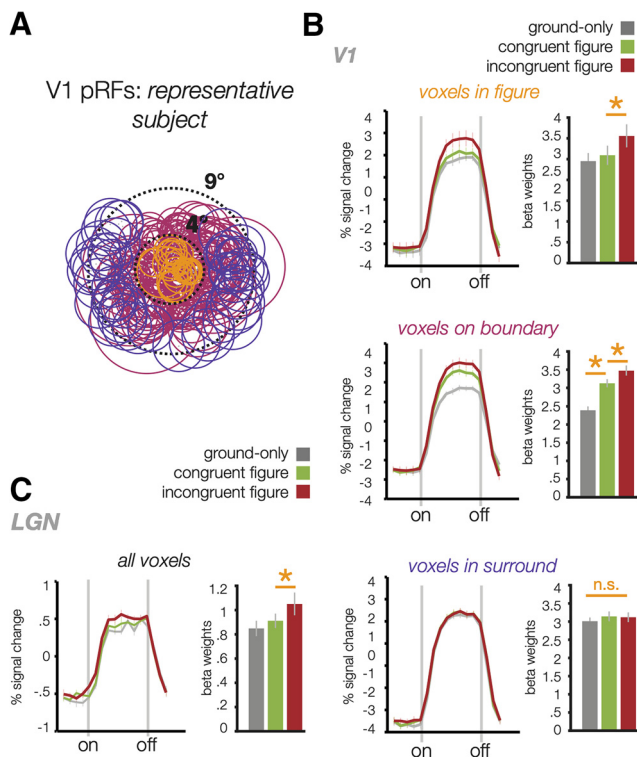
A

$$Y = X\beta + error$$

voxel responses = pRFs × spatial profile of effect + e

B

boundary response [C-G]   figure enhancement [I-C]   figure enhancement [I-C] (.5° gap)
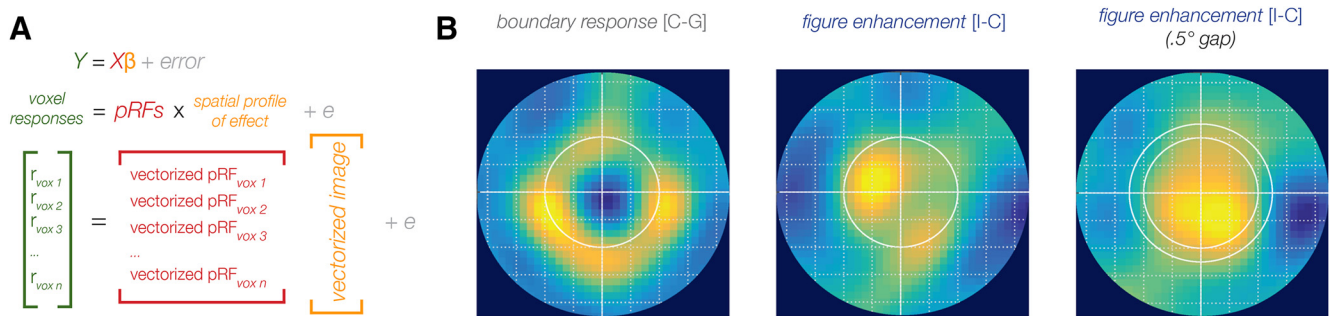
Figure 7. pRF-based visualization of voxelwise BOLD responses in stimulus space. ***A***, Schematic of the multivariate regression-based method of estimation, which allows us to estimate the spatial profile of the observed BOLD effect (yellow) from the measured pRFs (red) and BOLD response (green) of each voxel. The resulting estimate of the spatial profile (yellow) is in the form of a vector of length n^2, which is then reshaped to yield images of size $n \times n$ like those in ***B***. The precise formula used appears in the Materials and Methods. ***B***, The projected spatial profiles of the differential responses associated with boundary response (congruent minus ground-only condition, left) and figure enhancement (incongruent minus congruent) for V1 data in the main experiment (middle) and in a control experiment which introduced a 0.5° gap between figure and surround (right). Each pixel corresponds to 0.25° of visual angle of the original stimulus space; color depicts predictor values, normalized and averaged across subject. See also Figure 9.

that this enhancement is spatially distinct from responses associated with local processing of a phase-defined boundary. We ran an additional control experiment to ascertain whether the presence of a clearly visible boundary, rather than the orientation difference between figure and surround, might be sufficient to induce an effect of figure enhancement in V1. This control relied on the same stimulus configuration and task but introduced a 0.5° gap (2.0–2.5° eccentricity) between the figure and the surround (Fig. 8), so that the boundary between figure and surround could be very clearly perceived in both congruent and incongruent conditions. The reconstruction of V1 activity again shows that the orientation difference between figure and surround led to additionally enhanced responses throughout the extent of the central figure (Fig. 7B, right). Thus, we can conclude that figure enhancement cannot be explained simply in terms of an inward spread of local boundary responses; larger-scale differences in feature content that distinguish the figure from the surround confer an additional enhancement.

Behavioral performance and the role of attention
In recent work, we showed across several experiments that directed attention is not necessary to observe robust figure–ground modulation in the early visual cortex and in the LGN (Poltoratski et al., 2019). We did so first by requiring participants to attend to one of two bilaterally presented figures, and in another experiment by having participants perform an extremely challenging letter RSVP task at the central fixation. While these studies provided strong evidence that directed attention was not necessary for figure–ground modulation in early visual cortex, we were unable to resolve the relative contributions of local enhancement of the figure–surround boundary from that of the figure itself. This is of particular interest, since boundary detection is considered to occur preattentively, while it is debated whether figure enhancement is attenuated or extinguished in the absence of attention (Rossi et al., 2001; Marcus and van Essen, 2002; Poort et al., 2012). In the current experiments, experienced fMRI participants performed a sustained color change detection task on the fixation spot and were instructed that the figure–ground stimuli presented parafoveally would never be task relevant. Figure stimuli appeared in entirely predictable spatial positions with consistent timing throughout dozens of experimental blocks, and no participants reported difficulty in ignoring the figures to perform the task. Behavioral performance was unaffected by the temporal onset or offset of the figure stimuli and remained stable over time (Fig. 3A), implying that the
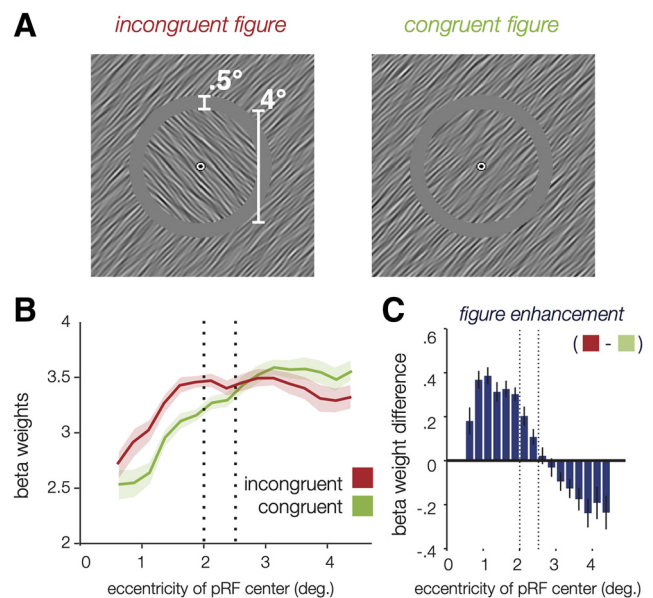

Figure 8. ***A–C***, Stimuli (***A***) and results (***B***, ***C***) of control experiment A, in which a 0.5° grayscale gap separated the figure and surround. Plotting follows conventions of Figure 2; here, since both conditions have a visible boundary, we can subtract responses of the congruent figure condition (green) from the incongruent figure condition (red) to yield orientation-dependent figure enhancement. Results closely follow the pattern of the main experiment, with the observed enhancement extending throughout the figure region.

appearance of the figures did not distract from performance of the central task. Additionally, we saw no significant differences in performance between any of the figure conditions relative to the ground-only condition, during which there was no figure present on the screen (Fig. 3B). Together, these results suggest that participants' attention was not systematically influenced by the task-irrelevant figures. We therefore attribute differential responses to the incongruent figure or congruent figure, relative to the ground-only condition, as indicative of automatic perceptual processes that do not require focal attention.

Results in areas V2 and V3
Our pRF mapping experiments indicated that pRF sizes were somewhat larger in V2 and V3 (Fig. 4A), as has been previously reported (Dumoulin and Wandell, 2008; Wandell and Winawer, 2015). Nevertheless, the pattern of results in V2 and V3 were very similar to those found in V1, in accord with recent

intracranial recordings of human V2/V3 neurons in response to feature-defined figures (Self et al., 2016). Note that the larger peripheral pRFs found in these extrastriate visual areas are more likely to encroach on the 2° figure–surround boundary, and thus border response effects are evident in a larger proportion of voxels than in V1. We again averaged the responses across voxels whose pRFs fell primarily in the figure, in the surround, or overlapped the boundary, assuming an FWHM central region. Voxels with pRFs confined well within the figure showed significant effects of figure enhancement (V2: $t_{(5)} = 5.37$, $p = 0.0030$; V3: $t_{(5)} = 4.81$, $p = 0.0048$) but did not show a significant boundary response (V2: $t_{(5)} = 1.38$, $p = 0.23$; V3: $t_{(5)} = 1.12$, $p = 0.31$). Voxels whose pRFs fell on the boundary exhibited both effects (V2: $t_{(5)} > 5.16$, $p < 0.0036$; V3: $t_{(5)} > 4.27$, $p < 0.0079$). In V2, voxels whose pRF envelopes were contained within the surround did not show evidence of either type of enhancement (V2: $t_{(5)} < 1.64$, $p > 0.16$). In V3, voxels corresponding to the surround also showed no effect of figure enhancement ($t_{(5)} = 0.80$, $p = 0.46$), though they did show a significant effect of boundary response ($t_{(5)} = 2.91$, $p = 0.034$). Of note, however, was that pRFs in this peripheral region of V3 were quite large and more spatially diffuse in their response profile; only 96 voxels across six participants were included in the surround-only subset of voxels.

Figure–ground modulation in the LGN

While our primary aim was to resolve the spatial extent of figure–ground modulation in the early visual cortex, we were also able to obtain measures of BOLD activity in the LGN. In recent work, we have shown that the LGN is sensitive to figure–ground modulation in the absence of directed attention (Poltoratski et al., 2019). Although the receptive fields of individual LGN neurons are small, each fMRI voxel covers a proportionally greater segment of this small subcortical structure than a similar voxel sampled from cortex. As such, the pRFs that we measured in the LGN were too large to clearly distinguish responses to the central figure region from those to the surround (eccentricity 2–4.5°); 96.2% of recorded voxels overlapped with the figure–surround boundary. However, we did find evidence of figure enhancement (incongruent > congruent figure: $t_{(5)} = 2.89$, $p = 0.034$) across the full set of LGN voxels, as shown in Figure 6C; there was no significant difference between responses in the congruent figure condition versus the ground-only condition ($t_{(5)} = 2.27$, $p = 0.072$).

We also noted a similar pattern of results in control experiment B, in which a larger figure was shown. When we analyzed the response of 50 voxels in this experiment whose pRFs fell on the figure region but also extended beyond the boundary, we observed significantly greater responses to incongruent figures than to congruent figures ($t_{(2)} = 8.2$, $p = 0.014$), consistent with our predicted effects of figure enhancement. In contrast, this LGN ROI did not show a significant difference in its response to congruent figures when compared with the ground-only condition ($t_{(2)} = 1.32$, $p = 0.32$), indicating a lack of sensitivity to the phase-defined border present in the congruent figure condition. We then performed a more spatially restricted analysis, focusing on the eight voxels whose pRFs fell within the figure region, and again found a significant effect of figure enhancement ($t_{(2)} = 4.7$, $p = 0.04$). These results suggest that activity in the lateral geniculate nucleus is involved in the broader visual circuitry that supports figure–ground perception, presumably via top–down feedback from the visual cortex. The current results serve to replicate and extend our previous findings (Poltoratski et al., 2019) and other recent work (Jones et al., 2015) describing the modulation of LGN responses by figure–ground configurations.

Discussion

Our study provides compelling evidence that early visual areas in the human brain are engaged in figure–ground processing; they respond to local boundary differences, but confer an additional enhancement of the entire figure region for figures that differ in featural content from their surround. We demonstrate that that this figure enhancement is spatially distinct from local responses associated with sensitivity to the boundary using a novel response reconstruction method informed by pRF mapping of voxels in the early visual cortex. While a phase-defined border was present in the congruent figure condition, this did not lead to widespread effects of figure enhancement, consistent with the weaker perceptual salience of the figure. It was also the case that introducing a clearly visible 0.5° gap around congruent-orientation figures was not sufficient to drive figure enhancement to the degree observed when the figure differed in orientation from the surround. Boundary enhancement is predicted by several known mechanisms, including a release from feature-tuned surround suppression that serves to enhance local orientation contrast (Blakemore and Tobin, 1972; Nelson and Frost, 1978; Nothdurft, 1991; Sillito et al., 1995; Bair et al., 2003; Shushruth et al., 2012) or sensitivity to high spatial frequency information at the boundary (Landy and Kojima, 2001; Mazer et al., 2002; Hallum et al., 2011). However, the asymmetric enhancement of the central figure region is not readily attributable to these well known neural mechanisms associated with local processing within V1.

Unlike boundary detection, which serves to enhance local feature differences and presumably involves local processing within V1 (Self et al., 2013; Bijanzadeh et al., 2018), figure enhancement appears to serve a more integrative function of grouping regions that share one or more features that distinguish it from the surround. Compelling neurophysiological work in the monkey has pointed to the role of feedback from higher cortical areas, including V4 (Self et al., 2012; Klink et al., 2017), as being critical for observing the effects of figure enhancement but not boundary detection in V1 (Lamme and Roelfsema, 2000). Enhanced responses to the border between figure and surround emerge rapidly in V1 neurons, and are evident in the initial transient response, whereas figure enhancement emerges later in time and remains sustained over longer viewing periods (Lamme and Roelfsema, 2000; Scholte et al., 2008; Poort et al., 2012). Accordingly, boundary detection and figure enhancement appear to modulate different laminar layers of V1, with the latter primarily leading to enhanced responses in layers 1, 2, and 5, which receive feedback from higher visual areas (Self et al., 2013). While fMRI does not provide information about the source of the modulations we observe here, our results are generally supportive of these proposed mechanisms of feedback. For voxels with pRFs that overlapped the figure, we found no relationship between the proximity of the response field of a voxel to the boundary and the magnitude of figure enhancement when contrasting responses to incongruent and congruent figures. Our findings deviate from standard accounts of orientation-tuned surround suppression, in which local inhibitory interactions lead to suppression that falls off symmetrically as a function of spatial separation between a target stimulus and the surround stimulus (Bair et al., 2003; Bijanzadeh et al., 2018). They also run counter to alternative theories that challenge the need for mechanisms of figure enhancement beyond detection of the boundary (Rossi et al., 2001; Zhaoping, 2003). Previous studies were not able to characterize the spatial profile of these distinct mechanisms in humans as we were able to carry out here using high-field
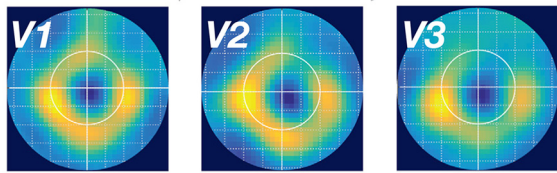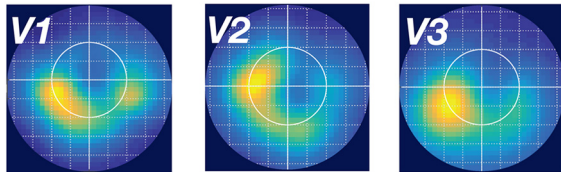
regression-based projection

weighted pRF average

Figure 9. A comparison of two methods for visualization of the boundary response (congruent minus baseline) in areas V1, V2, and V3. The top panel shows our ridge regression-based approach, while the bottom panel is a simpler average of each pRF weighted by the BOLD amplitude of boundary response-associated BOLD beta weight difference in that voxel. The regression-based approach generally led to better resolved reconstructions, particularly in higher visual areas, which typically have larger pRFs.

imaging, pRF modeling, and computational methods for regression-based visualization.

To date, figure processing in humans has been largely studied in higher visual areas, including hV4 (Kastner et al., 2000; Kourtzi et al., 2003; Thielscher et al., 2008) and the lateral occipital cortex (Grill-Spector et al., 1999; Vinberg and Grill-Spector, 2008); indeed, some early neuroimaging work did not find modulations in V1 in response to feature-defined textures or figures, likely due to insufficient signal strength or sensitivity (Kastner et al., 2000; Schira et al., 2004). However, this work and that of others (Appelbaum et al., 2008; Scholte et al., 2008; Self et al., 2016; Poltoratski et al., 2019) provide convergent evidence of common mechanisms of figure–ground processing across human and nonhuman primates, in which processing in higher-level visual areas including V4 informs, via feedback, enhancement of responses to figure regions in V1 and the LGN.

Performing this work in humans allows us to leverage several methodological advantages and to build a clearer understanding of the role of attention in figure–ground modulation. In particular, the 2D reconstruction of figure–ground modulation (Fig. 7) provides compelling evidence of the spatial profile of its constituent mechanisms. fMRI uniquely allows us to simultaneously monitor the activity of thousands of voxels in the early visual system, and expands our knowledge of how populations of neurons in early visual regions encode perceptual figures. Our results provide satisfying convergence with neurophysiological recordings in the monkey visual cortex (Lamme, 1995; Zipser et al., 1996; Self et al., 2019) while allowing us to manipulate participants' task and locus of attention easily and flexibly, without the need for reward-based training. This is particularly important since reward-based training, such as that used in the majority of neurophysiology studies, has been shown to alter the visual salience of trained stimuli for significant periods of time (Anderson et al., 2011; Anderson and Yantis, 2013).

In building convergence between human and nonhuman primate methods, this work extends several recent findings of V1 contributions to mechanisms of visual segmentation and grouping, including visual salience (Zhang et al., 2012; Poltoratski et al., 2017), grouping (Murray et al., 2002; Kourtzi et al., 2003;

Roelfsema, 2006), perceptual filling-in (Sasaki and Watanabe, 2004; Meng et al., 2005; Hong and Tong, 2017), and the processing of illusory surfaces (Kok and de Lange, 2014; Kuai et al., 2017). Together, these studies point to an important role of the early visual system in figure perception through a combination of both local feedforward mechanisms and automatic feedback mechanisms, which provide contextually driven modulation of responses in the absence of directed attention or particular task demands. While it may be the case that attention differentially modulates figure enhancement and boundary detection (Lamme et al., 1998b; Qiu et al., 2007; Poort et al., 2012; Self et al., 2019), consistent with their separable mechanistic origins, growing evidence suggests that figure enhancement is not a simple by-product of spatial attention (Marcus and van Essen, 2002; Jones et al., 2015; Papale et al., 2018; Poltoratski et al., 2019). Indeed, it appears that both humans and monkeys exhibit spatially specific effects of figure enhancement irrespective of the locus of attention, task, or training exposure to feature-defined figures (Poltoratski et al., 2019; Self et al., 2019).

This study also provides the second reported evidence of figure enhancement in human LGN (Poltoratski et al., 2019), which highlights an expanding role for this thalamic region in the complex visual processing of perceptual figures (Briggs and Usrey, 2008; Ghodrati et al., 2017). As figure–ground modulation is thought to rely on form-selective and object-selective regions beyond primary visual cortex, this implies a robust transfer of information from higher-order visual areas to the earliest possible anatomic stage of the visual system that can be modulated by neural feedback. This is also particularly interesting because LGN neurons themselves exhibit only modest orientation tuning (Leventhal and Schall, 1983; Xu et al., 2002) and untuned extraclassical RF suppression (Alitto and Usrey, 2008); however, via feedback, LGN responses can be sensitive not only to orientation information, but also to second-order feature contrast. Widespread cortical feedback to the LGN provides considerable opportunity for sculpting and refining the responses that are carried forward by the LGN (Wang et al., 2006; Briggs and Usrey, 2011). Feedback projections to the LGN reveal a high degree of retinotopic specificity, which provides a viable computational architecture for the spatial enhancement of orientation-defined figures.

Finally, this study highlights a useful and intuitive method for using population-receptive field modeling to reconstruct visual responses (Thirion et al., 2006; Miyawaki et al., 2008; Kok and de Lange, 2014), which can move toward bridging the gap between the resolution of spatial effects measured by neurophysiological recordings and human neuroimaging. The regression method allows for a finer spatial resolution of reconstruction than a simple summation of pRF responses weighted by their response amplitudes (Fig. 9). The latter approach is necessarily constrained by the resolution of pRFs themselves: even perfectly noise-free data will yield blurry reconstructions if the Gaussian pRFs are large, relative to the stimulus, and greater effects of pRF blurring will occur when ascending the visual hierarchy. This regression-based reconstruction approach could be adopted more widely to estimate the spatial profile of fMRI BOLD effects at spatial scales that would prove challenging with standard approaches.

References

Aghajari S, Vinke LN, Ling S (2020) Population spatial frequency tuning in human early visual cortex. J Neurophysiol 123:773–785.

Alitto HJ, Usrey WM (2008) Origin and dynamics of extraclassical suppression in the lateral geniculate nucleus of the macaque monkey. Neuron 57:135–146.

Anderson BA, Yantis S (2013) Persistence of value-driven attentional capture. J Exp Psychol Hum Percept Perform 39:6–9.

Anderson BA, Laurent PA, Yantis S (2011) Value-driven attentional capture. Proc Natl Acad Sci U S A 108:10367–10371.

Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM (2006) Cue-invariant networks for figure and background processing in human visual cortex. J Neurosci 26:11695–11708.

Appelbaum LG, Wade AR, Pettet MW, Vildavski VY, Norcia AM (2008) Figure-ground interaction in the human visual cortex. J Vis 8(9):8, 1–19.

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.

Bergen JR, Adelson EH (1988) Early vision and texture perception. Nature 333:363–364.

Bijanzadeh M, Nurminen L, Merlin S, Clark AM, Angelucci A (2018) Distinct laminar processing of local and global context in primate primary visual cortex. Neuron 100:259–274.e4.

Blakemore C, Tobin EA (1972) Lateral inhibition between orientation detectors in the cat's visual cortex. Exp Brain Res 15:439–440.

Briggs F, Usrey WM (2008) Emerging views of corticothalamic function. Curr Opin Neurobiol 18:403–407.

Briggs F, Usrey WM (2011) Corticogeniculate feedback and visual processing in the primate. J Physiol 589:33–40.

Cohen EH, Tong F (2015) Neural mechanisms of object-based attention. Cereb Cortex 25:1080–1092.

Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. Neuroimage 39:647–660.

Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7:181–192.

Ghodrati M, Khaligh-Razavi S-M, Lehky SR (2017) Towards building a more complex view of the lateral geniculate nucleus: recent advances in understanding its role. Prog Neurobiol 156:214–255.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron 24:187–203.

Hallum LE, Landy MS, Heeger DJ (2011) Human primary visual cortex (V1) is selective for second-order spatial frequency. J Neurophysiol 105:2121–2131.

Hastings C Jr, Mosteller F, Tukey JW (1947) Low moments for small samples: a comparative study of order statistics. Ann Math Statist 8:413–426.

Henriksson L, Nurminen L, Hyvärinen A, Vanni S (2008) Spatial frequency tuning in human retinotopic visual areas. J Vis 8(10):5, 1–13.

Himmelberg MM, Wade AR (2019) Eccentricity-dependent temporal contrast tuning in human visual cortex measured with fMRI. Neuroimage 184:462–474.

Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. J Technometrics 12:55–67.

Hong SW, Tong F (2017) Neural representation of form-contingent color filling-in in the early visual cortex. J Vis 17(13):10, 1–10.

Johnston A, Wright MJ (1985) Lower thresholds of motion for gratings as a function of eccentricity and contrast. Vision Res 25:179–185.

Jones HE, Andolina IM, Shipp SD, Adams DL, Cudeiro J, Salt TE, Sillito AM (2015) Figure-ground modulation in awake primate thalamus. Proc Natl Acad Sci U S A 112:7085–7090.

Kastner S, De Weerd P, Ungerleider LG (2000) Texture segregation in the human visual cortex: a functional MRI study. J Neurophysiol 83:2453–2457.

Klink PC, Dagnino B, Gariel-Mathis M-A, Roelfsema PR (2017) Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. Neuron 95:209–220.e3.

Kok P, de Lange FP (2014) Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. Curr Biol 24:1531–1535.

Kourtzi Z, Tolias AS, Altmann CF, Augath M, Logothetis NK (2003) Integration of local features into global shapes. Neuron 37:333–346.

Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60:1126–1141.

Kuai S-G, Li W, Yu C, Kourtzi Z (2017) Contour integration over time: psychophysical and fMRI evidence. Cereb Cortex 27:3042–3051.

Kwon O-S, Tadin D, Knill DC (2015) Unifying account of visual motion and position perception. Proc Natl Acad Sci U S A 112:8142–8147.

Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 15:1605–1615.

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

Lamme VA, Super H, Spekreijse H (1998a) Feedforward, horizontal, and feedback processing in the visual cortex. Curr Opin Neurobiol 8:529–535.

Lamme VA, Zipser K, Spekreijse H (1998b) Figure-ground activity in primary visual cortex is suppressed by anesthesia. Proc Natl Acad Sci U S A 95:3263–3268.

Landy MS, Bergen JR (1991) Texture segregation and orientation gradient. Vision Res 31:679–691.

Landy MS, Kojima H (2001) Ideal cue combination for localizing texture-defined edges. J Opt Soc Am A 18:2307–2320.

Leventhal AG, Schall JD (1983) Structural basis of orientation sensitivity of cat retinal ganglion cells. J Comp Neurol 220:465–475.

Marcus DS, van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. J Neurophysiol 88:2648–2658.

Mazer JA, Vinje WE, McDermott J, Schiller PH, Gallant JL (2002) Spatial frequency and orientation tuning dynamics in area V1. Proc Natl Acad Sci U S A 99:1645–1650.

Meng M, Remus DA, Tong F (2005) Filling-in of visual phantoms in the human brain. Nat Neurosci 8:1248–1254.

Miyawaki Y, Uchida H, Yamashita O, Sato M-A, Morito Y, Tanabe HC, Sadato N, Kamitani Y (2008) Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron 60:915–929.

Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL (2002) Shape perception reduces activity in human primary visual cortex. Proc Natl Acad Sci U S A 99:15164–15169.

Nelson JI, Frost BJ (1978) Orientation-selective inhibition from beyond the classic visual receptive field. Brain Res 139:359–365.

Nothdurft HC (1991) Texture segmentation and pop-out from orientation contrast. Vision Res 31:1073–1078.

Nothdurft HC (1993) The role of features in preattentive vision: comparison of orientation, motion and color cues. Vision Res 33:1937–1958.

Papale P, Leo A, Cecchetti L, Handjaras G, Kay KN, Pietrini P, Ricciardi E (2018) Foreground-background segmentation revealed during natural image viewing. eNeuro 5:ENEURO.0075-18.2018–18.

Poltoratski S, Ling S, McCormack D, Tong F (2017) Characterizing the effects of feature salience and top-down attention in the early visual system. J Neurophysiol 118:564–573.

Poltoratski S, Maier A, Newton AT, Tong F (2019) Figure-ground modulation in the human lateral geniculate nucleus is distinguishable from top-down attention. Curr Biol 29:2051–2057.e3.

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.

Poort J, Self MW, van Vugt B, Malkki H, Roelfsema PR (2016) Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex. Cereb Cortex 26:3964–3976.

Qiu FT, Sugihara T, von der Heydt R (2007) Figure-ground mechanisms provide structure for selective attention. Nat Neurosci 10:1492–1499.

Roelfsema PR (2006) Cortical algorithms for perceptual grouping. Annu Rev Neurosci 29:203–227.

Rossi A, Desimone R, Ungerleider L (2001) Contextual modulation in primary visual cortex of macaques. J Neurosci 21:1698–1709.

Sasaki Y, Watanabe T (2004) The primary visual cortex fills in color. Proc Natl Acad Sci U S A 101:18251–18256.

Schira MM, Fahle M, Donner TH, Kraft A, Brandt SA (2004) Differential contribution of early visual areas to the perceptual process of contour processing. J Neurophysiol 91:1716–1721.

Scholte HS, Jolij J, Fahrenfort JJ, Lamme VAF (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 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.

Self MW, Peters JC, Possel JK, Reithler J, Goebel R, Ris P, Jeurissen D, Reddy L, Claus S, Baayen JC, Roelfsema PR (2016) The effects of context and attention on spiking activity in human early visual cortex. PLoS Biol 14: e1002420.

Self MW, Jeurissen D, van Ham AF, van Vugt B, Poort J, Roelfsema PR (2019) The segmentation of proto-objects in the monkey primary visual cortex. Curr Biol 29:1019–1029.e4.

Shushruth S, Mangapathy P, Ichida JM, Bressloff PC, Schwabe L, Angelucci A (2012) Strong recurrent networks compute the orientation tuning of surround modulation in the primate primary visual cortex. J Neurosci 32:308–321.

Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J (1995) Visual cortical mechanisms detecting focal orientation discontinuities. Nature 378:492–496.

Singh KD, Smith AT, Greenlee MW (2000) Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. Neuroimage 12:550–564.

Somers DC, Dale AM, Seiffert AE, Tootell RB (1999) Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. Proc Natl Acad Sci U S A 96:1663–1668.

Supèr H, Spekreijse H, Lamme VA (2001) Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). Nat Neurosci 4:304–310.

Swisher JD, Sexton JA, Gatenby JC, Gore JC, Tong F (2012) Multishot versus single-shot pulse sequences in very high field fMRI: a comparison using retinotopic mapping. PLoS One 7:e34626.

Thielscher A, Kölle M, Neumann H, Spitzer M, Grön G (2008) Texture segmentation in human perception: a combined modeling and fMRI study. Neuroscience 151:730–736.

Thirion B, Duchesnay E, Hubbard E, Dubois J, Poline J-B, Lebihan D, Dehaene S (2006) Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage 33:1104–1116.

Tootell RB, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM (1998) The retinotopy of visual spatial attention. Neuron 21:1409–1422.

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12:97–136.

Vinberg J, Grill-Spector K (2008) Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. J Neurophysiol 99:1380–1393.

Wandell BA, Winawer J (2015) Computational neuroimaging and population receptive fields. Trends Cogn Sci 19:349–357.

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366–383.

Wang W, Jones HE, Andolina IM, Salt TE, Sillito AM (2006) Functional alignment of feedback effects from visual cortex to thalamus. Nat Neurosci 9:1330–1336.

Xu X, Ichida J, Shostak Y, Bonds AB, Casagrande VA (2002) Are primate lateral geniculate nucleus (LGN) cells really sensitive to orientation or direction? Vis Neurosci 19:97–108.

Zhang X, Zhaoping L, Zhou T, Fang F (2012) Neural activities in V1 create a bottom-up saliency map. Neuron 73:183–192.

Zhaoping L (2003) V1 mechanisms and some figure–ground and border effects. J Physiol Paris 97:503–515.

Zipser K, Lamme VA, Schiller PH (1996) Contextual modulation in primary visual cortex. J Neurosci 16:7376–7389.