NEUROPHYSIOLOGY

Contextual drive of neuronal responses in mouse V1 in the absence of feedforward input

Lisa Kirchberger[1], Sreedeep Mukherjee[1], Matthew W. Self[1]†, Pieter R. Roelfsema[1,2,3,4]*†

Neurons in the primary visual cortex (V1) respond to stimuli in their receptive field (RF), which is defined by the feedforward input from the retina. However, V1 neurons are also sensitive to contextual information outside their RF, even if the RF itself is unstimulated. Here, we examined the cortical circuits for V1 contextual responses to gray disks superimposed on different backgrounds. Contextual responses began late and were strongest in the feedback-recipient layers of V1. They differed between the three main classes of inhibitory neurons, with particularly strong contextual drive of VIP neurons, indicating a contribution of disinhibitory circuits to contextual drive. Contextual drive was strongest when the gray disk was perceived as figure, occluding its background, rather than a hole. Our results link contextual drive in V1 to perceptual organization and provide previously unknown insight into how recurrent processing shapes the response of sensory neurons to facilitate figure perception.

INTRODUCTION

Neurons in primary visual cortex (V1) are often conceptualized as simple feature detectors, signaling the presence of visual stimuli in a localized area of visual space, their receptive field (RF). However, the activity of V1 neurons also depends on the context determined by visual information outside their RF. In most conditions, the visual context has a modulatory influence on the response elicited by the RF stimulus (*1*–*3*). These modulatory influences are thought to depend on recurrent connections, i.e., horizontal connections within V1 (*4*, *5*) and feedback connections from higher visual areas (*6*–*10*). The modulatory influences on neuronal responses are typically absent during the initial visual response elicited by the onset of a stimulus in the RF and are frequently observed after a delay, likely caused by synaptic delays associated with the longer, recurrent processing routes. A prototypical contextual effect observed in V1 is figure-ground modulation (*11*). Studies in the visual cortex of humans (*12*), monkeys (*11*), and mice (*13*) found that neural activity is enhanced when the RF falls on an area of the visual field that is part of a figure, compared to the background. Figure-ground modulation even occurs if the visual information within the RF is held constant and it depends on feedback from higher visual areas (*14*).

While most previous studies focused on the contextual modulation of responses elicited by stimuli located within and beyond the RF, there are also conditions in which the context alone drives the activity of V1 neurons, even if their RF is not stimulated. For example, some V1 neurons are driven by illusory contours (Fig. 1A) (*15*–*17*). Another condition in which V1 neurons have been shown to be activated without a stimulus in the RF is when the RF falls on a uniform, blank disk embedded in a larger background (Fig. 1B). This phenomenon was first described in monkeys by Lee *et al.* in 1998 (*18*), who related it to figure-ground perception, and it was later also observed in mice (*13*). In a recent study, Keller *et al.* (*19*) reported that these contextually driven responses depend on feedback from higher cortical areas and proposed that V1 neurons have a second "feedback RF" (fbRF). Keller *et al.* (*19*) hypothesized that the fbRF signals differences between the surface in the RF and other image regions, in accordance with theories of predictive coding (*20*, *21*). A third phenomenon that may explain the origin of the contextually evoked activity in V1, could be "filling-in" of occluded image regions (*22*), similarly to filling in the retinal blind spot with information from the surrounding areas. In this study, we characterized the laminar and cell type–specific profile of V1 responses evoked by contextual stimuli lacking classical RF stimulation and asked whether the contextually evoked activity depends on figure-ground organization or may be best explained by filling in.

We presented a gray disk in the RF of V1 neurons that was surrounded by a grating and observed robust neuronal responses, which were strongest in feedback-recipient layers. The luminance of the gray disk was identical to that of the gray screen during the intertrial interval (ITI) so that stimulus onset was not associated with any change in the part of the visual stimulus that fell within the RF. To gain mechanistic insight, we used cell-specific Ca^{2+} imaging to examine the responses of layer 2/3 pyramidal neurons and the three main classes of interneurons: parvalbumin-positive (PV), somatostatin-expressing (SST), and vasoactive intestinal peptide–expressing (VIP) interneurons. The cell types were differentially activated by contextual stimuli, and we examined the possible role of a disinhibitory circuit involving VIP neurons in contextual activation of V1 neurons (*23*–*27*). We also asked whether contextual responses are related to figure-ground organization by testing whether the strength of figure-ground modulation predicts the strength of the contextual response and whether it matters if the disk in the RF is perceived as part of a figure or the background. We report that V1 activity was consistently enhanced for disks that are perceived as figures, indicating that the contextual drive may be closely linked to perceptual organization.

[1]Department of Vision and Cognition, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, Netherlands. [2]Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU University, Amsterdam, Netherlands. [3]Department of Psychiatry, Academic Medical Center, Amsterdam, Netherlands. [4]Laboratory of Visual Brain Therapy, Sorbonne Université, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Institut de la Vision, Paris F-75012, France.
*Corresponding author. Email: p.roelfsema@nin.knaw.nl
†These authors contributed equally to this work as last authors.

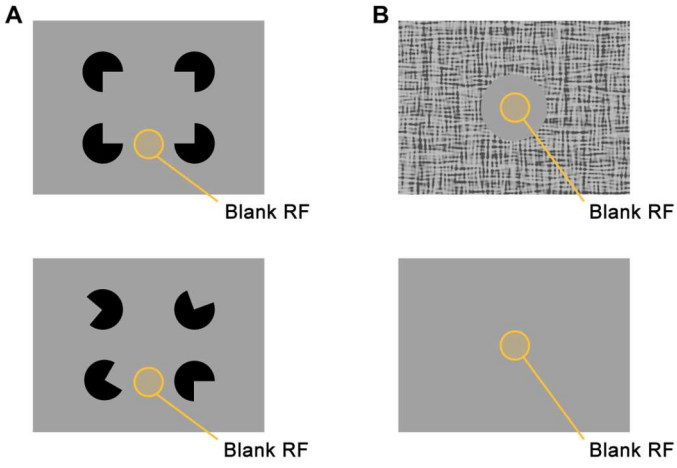

Fig. 1. Stimuli evoking V1 contextual activation without stimulating the RF. (**A**) Top: An illusory figure is evoked by the alignment of the pacmen, which is absent when the pacmen are rotated (bottom). (**B**) Schematic of a blank circular image region embedded in a larger background, evoking activity even if the RF falls in the center of the circle.

RESULTS

V1 activity elicited by visual stimuli outside the classical RF

We first examined the properties of contextual drive in V1. We presented a grating stimulus in a 40° circular aperture ($RF_{grating}$ stimulus; Fig. 2B) to nine awake mice that could run on a treadmill (on average 60 repeats, spatial frequency of 0.075 cycles/degree, and mean luminance of 20 cd/m^2). On other trials, we presented a full-screen grating (uniform stimulus; Fig. 2B) or a full-screen grating with a 40° homogenous gray disk centered on, and fully containing, the RF of the recorded neurons (RF_{blank} stimulus; Fig. 2B; see Materials and Methods for details of all stimuli). For the RF_{blank} stimulus, the luminance of the gray disk was equal to that of the gray screen during the ITI, such that no change occurred in the RF of the recorded neurons upon presentation of the stimulus. We recorded extracellular multi-unit activity using 32-channel laminar silicon probes inserted in V1. We first mapped the RF location of the recorded neurons using sparse noise stimuli (Fig. 2D, left for an example) and centered the 40° gratings and disks on that location. We applied strict inclusion criteria to ensure that the recorded neurons were not driven by the boundary between the gray disk and the surround grating, or by the surround itself, by only including neurons with RFs fully contained within the gray disk [see Materials and Methods and Fig. 2D (right)]. This led to the inclusion of 401 recording sites across 27 electrode penetrations in nine mice. We normalized the responses at individual recording sites to the peak response elicited by the full-screen grating.

In accordance with previous studies (*13*, *18*, *19*), we observed that all visual stimuli elicited robust V1 activity (panels B-D of Fig. 2 show an example site, and Fig. 2E shows the population).

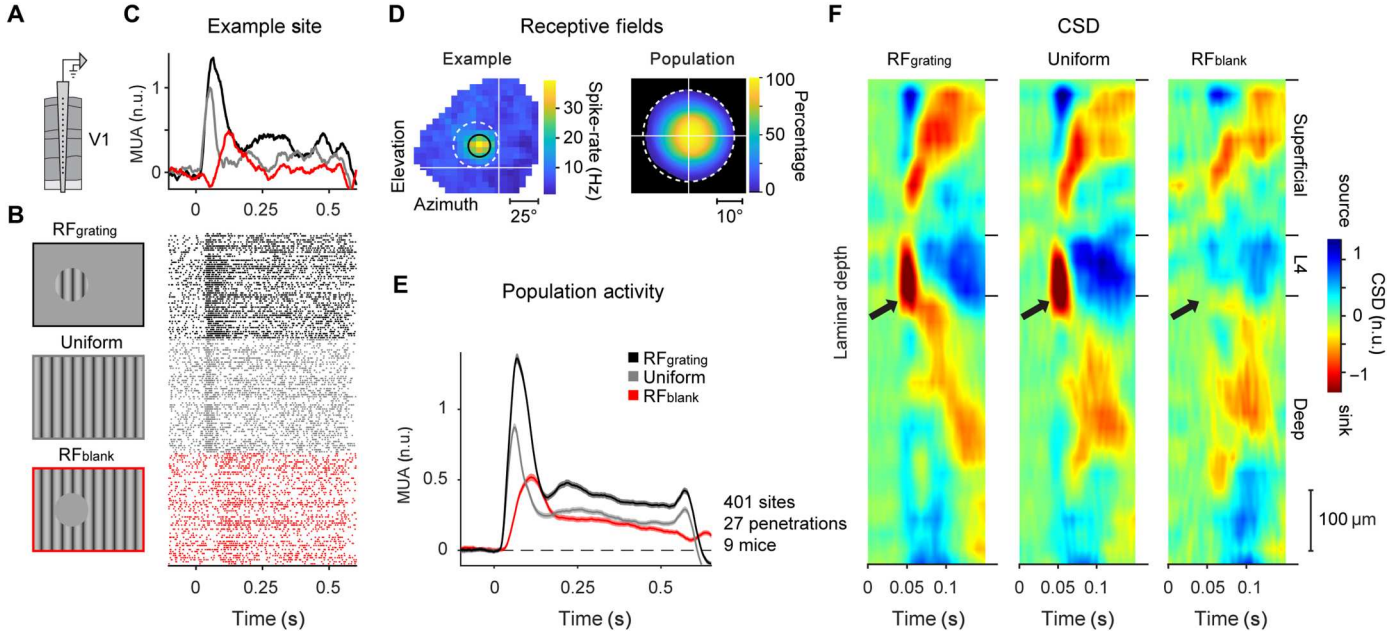

Fig. 2. V1 responses to gray disk stimuli. (**A**) We used a 32-channel laminar silicon probe to record V1 activity. n.u., normalized units. (**B**) We presented visual stimuli with a circular grating centered on the RF ($RF_{grating}$ stimulus, top, black), a uniform, full-screen grating (uniform stimulus, middle, gray) or a gray disk occluding a region of a full-screen grating, which was centered on the RF of the recorded neurons (RF_{blank}, bottom, red). (**C**) Top: Average activity of neurons at the example site located in the superficial layers. Bottom: Raster plots of activity at this recording site. (**D**) Left: RF of the example recording site presented in (C). Black circle, Gaussian fit to the RF. White dashed circle, figure position. Right: RF coverage in percent relative to the center of the $RF_{grating}$/RF_{blank} stimulus. We excluded all units with RFs extending beyond the edge (white dashed circle). (**E**) Activity evoked by the three stimuli across a population of 401 recording sites in 27 electrode penetrations in nine mice. (**F**) Average, normalized CSD profile across all 27 penetrations evoked by the $RF_{grating}$ (left), uniform (middle), and RF_{blank} (right) stimulus.

The response to the RF_{blank} stimulus occurred despite the absence of a change in the RF when the stimulus appeared. A linear mixed-effects (LME) model, which accounts for the hierarchical correlation structure in the data with different penetrations and mice, revealed that the contextual response was significant ($P < 0.001$; see Materials and Methods for details). To assign the recording sites to the different layers, we determined the current source density (CSD) profile for each penetration and localized the boundary between layers 4 and 5 (*28*, *29*). We then aligned the recording sites of all penetrations to this L4/5 border and calculated the average CSD to examine the laminar profile of current sinks and sources evoked by the three visual stimuli (Fig. 2F). The stimuli with gratings stimulating the RF elicited the typical, prominent current sink in layer 4 (black arrows in Fig. 2F). This current sink was absent for the RF_{blank} stimulus, confirming that the V1 response was not driven by edge or grating influences via feedforward input from lateral geniculate nucleus (LGN). Instead, the responses to this stimulus were likely evoked by recurrent connections, i.e., horizontal connections within V1 and/or feedback connections from higher brain areas (*19*).

Activity elicited by the gray disk is delayed and stronger in feedback-recipient layers

While feedforward connections in the neocortex preferentially target layer 4, feedback connections tend to avoid layer 4 and innervate the superficial and deep layers (*30*–*33*). We characterized the laminar profiles of V1 activity elicited by the gray disk and grating stimuli (Fig. 3A). We assigned all recording sites to one of three laminar compartments: superficial (100 to 400 μm above L4/5 border), layer 4 (0 to 100 μm above L4/5 border), and deep (0 to 600 μm below L4/5 border). We determined the latency of the visual responses for each recording site by fitting a curve to the activity time courses and defined the latency as the point at which the curve reached 33% of its maximum (see Materials and Methods; Fig. 3D) (*34*). The latency of the visual response evoked by the $RF_{grating}$ and uniform stimulus was around 40 ms and shortest in layer 4 ($RF_{grating}$: 45.2 ± 0.5 ms in the superficial layers, 40.8 ± 0.4 ms in L4, and 46 ± 0.5 ms in the deep layers; $P < 0.001$ for L4 versus superficial or deep and $P < 0.05$ for superficial versus deep, LME; uniform: 43.1 ± 2.2 ms in the superficial layers, 39.2 ± 0.6 ms in L4, and 46.5 ± 1.0 ms in the deep layers; $P < 0.01$ for L4 versus deep and not significant for other comparisons; Fig. 3E).

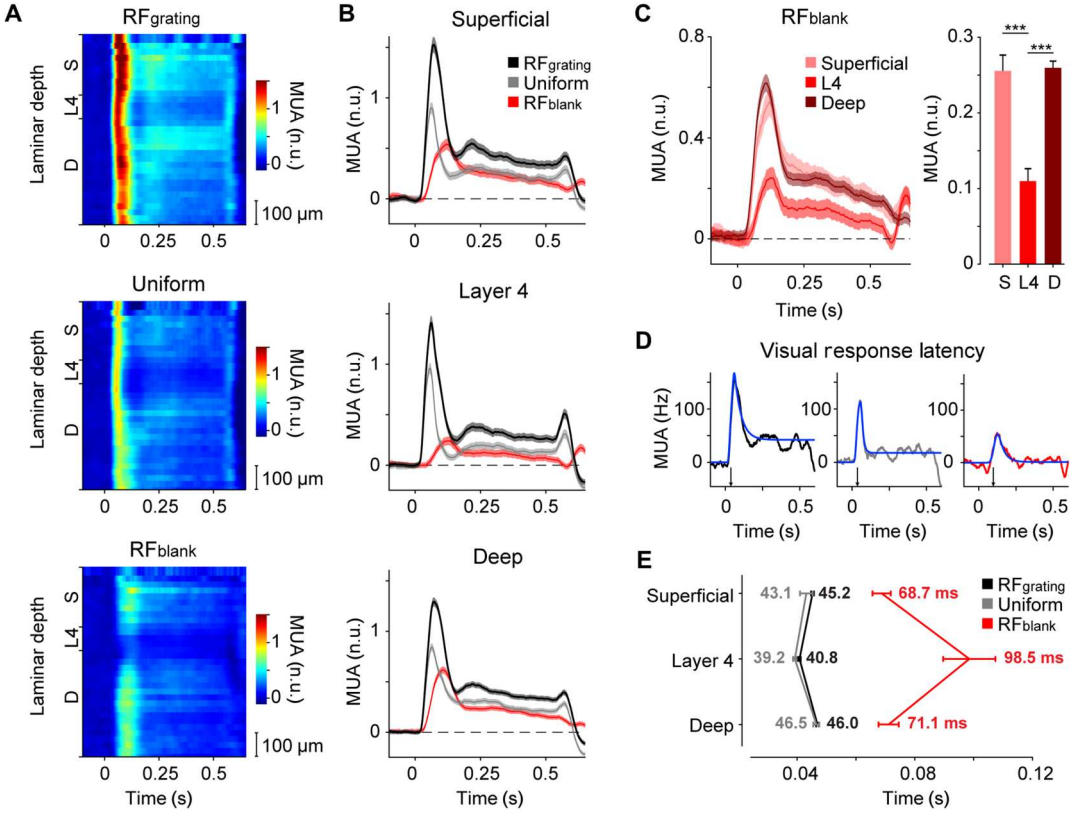

Fig. 3. Neuronal responses to gray disk stimuli are delayed and strongest in the superficial and deep layers. (**A**) Average V1 response across the cortical depth (25-μm distance between recording sites) evoked by the $RF_{grating}$ (top), uniform (middle), and RF_{blank} (bottom) stimulus ($n = 401$ recording sites from 27 electrode penetrations in nine mice). (**B**) Average V1 response in the superficial layers ($n = 115$ sites, top), layer 4 ($n = 74$, middle), and deep layers ($n = 212$, bottom) evoked by the $RF_{grating}$ (black), uniform (gray), or RF_{blank} (red) stimulus. (**C**) Left: Time course of activity evoked by the RF_{blank} stimulus for the superficial (light pink), L4 (red), and deep (dark red) recording sites. Right: Activity evoked by the RF_{blank} stimulus from 0 to 0.5 s after stimulus onset (normalized to the peak response elicited by the uniform grating; see Materials and Methods). Activity in the superficial and deep layers was stronger than that in layer 4 ($P < 0.001$, LME). (**D**) We fit curves to the time course of this example recording site to determine the visual response latency (blue). (**E**) Average visual response latency in the three laminar compartments evoked by the $RF_{grating}$ (black), uniform (gray), or RF_{blank} (red) stimulus. The onset of the response to the RF_{blank} stimulus was delayed compared to that to the $RF_{grating}$ and uniform stimuli, in all laminar compartments ($P < 0.001$, LME).

The short latency in layer 4 is in accordance with the relatively strong feedforward inputs from the LGN in this layer. Layer 4 neurons then activate neurons in superficial and deep layers, and the longer latency in the other layers can be explained by the additional synaptic delays. In contrast, the RF_{blank} stimulus lacks visual information in the RF, and there is no feedforward input from the LGN. Accordingly, visual onset latencies were longer for the RF_{blank} stimulus than for the stimuli with a grating in the RF, across all layers ($P < 0.001$, LME). The pattern of latencies across the layers was inverted with the shortest latencies in the superficial and deep layers and the longest latencies in layer 4 (68.7 ± 3.1 ms in the superficial layers, 98.5 ± 8.9 ms in L4, and 71.1 ± 3.4 ms in the deep layers; $P < 0.05$ for L4 versus superficial or deep, LME; Fig. 3E). The longer latencies in layer 4 for the gray disk stimulus were accompanied by a weaker response (Fig. 3C), which was significantly lower than in the feedback-recipient superficial and deep layers ($P < 0.001$, LME, for comparison between superficial and layer 4, and between deep and layer 4; Fig. 3, A and C). Together, these results demonstrate that visual stimuli outside the classical RF can evoke robust responses in the superficial and deep layers of V1, in accordance with previous work (*13*, *19*).

Responses of excitatory and different types of inhibitory neurons to gray disk stimuli

To gain mechanistic insight into the V1 microcircuit involved in processing contextual information, we examined responses of layer 2/3 pyramidal neurons and the activity of three main classes of cortical interneurons using cell-specific two-photon imaging of Ca^{2+} activity. We measured the activity of pyramidal neurons by implanting five Thy1-GCaMP6f animals (*35*) with cranial windows and head fixation rings and additionally injected AAV1-CaMKII-GCaMP6f-WPRE-SV40 in V1 to increase the signal-to-noise ratio (SNR) of the Ca^{2+} signals of neurons in layer 2/3. We placed the mice on a treadmill, determined the RFs of the recorded V1 neurons, and measured the activity elicited by the different stimuli. Just as in the electrophysiological experiments, the gray disk stimuli elicited robust activity in layer 2/3 (Fig. 4B). The average response across 537 excitatory neurons evoked by the RF_{blank} stimulus was 58.2% of the response evoked by a full-screen grating (Fig. 4, B and C) a difference that was significant ($P < 0.001$, LME). Furthermore, the majority of cells gave a stronger response to the full-screen grating than to the gray disk (325 of 537 cells or 60.5%).

We also measured the responses of PV, SST, and VIP interneurons. The activity elicited by the RF_{blank} stimulus in PV cells ($n = 87$ cells in four mice), which are thought to control the gain of the cortical column, was comparable to that elicited by the full-screen grating (time windows of 0 to 2 s after stimulus onset, $P = 0.97$, LME; Fig. 4D). The time course of the response differed between stimuli. The initial response of PV neurons was stronger for full-screen stimuli, but the response elicited by the gray disk stimulus became stronger after 1 s. VIP neurons in V1 are among the targets of feedback projections (*24*, *27*) and inhibit SST neurons (*27*), which, in turn, inhibit the apical dendrites of pyramidal neurons. Through this disinhibitory motif VIP neurons have a net excitatory effect on pyramidal neurons. VIP neurons in V1 were more strongly driven by the RF_{blank} stimulus than by full-screen stimuli ($n = 120$ in four mice; $P < 0.05$, LME; Fig. 4E). SST neurons exhibited a strong response to the full-screen grating. They were slightly inhibited by the RF_{blank} stimulus ($n = 70$ SST neurons in five mice) so that this response was much weaker than that evoked by the full-screen grating ($P < 0.001$, LME; Fig. 4F). These activity patterns of VIP and SST neurons are in accordance with Keller *et al.* (*19*) and are also in line with the disinhibitory circuit, in which VIP neurons inhibit SST-neurons, contributing to the gray disk response.

Involvement of the VIP-mediated disinhibitory circuit in the gray disk response

To directly test the involvement of the VIP neuron–mediated disinhibitory circuit in the gray disk response, we used electrophysiology to record V1 activity while optogenetically inhibiting the activity of VIP neurons. We expressed the inhibitory opsin GtACR2 in VIP neurons in V1 of eight VIP-Cre mice with a Cre-dependent viral vector (AAV-hSyn-SIO-stGtACR2-FusionRed; Fig. 5A). This led to the expression of the soma-targeted version of the chloride channel GtACR2 in VIP neurons in V1, such that exposure to blue light inhibited the activity of VIP neurons. Inhibition of VIP neurons reduced the response to the gray disk by 26% ($P < 0.001$, LME; Fig. 5, B and C), an effect that occurred across all layers ($P < 0.001$ for all depths, LME; fig. S1A). Furthermore, inhibiting VIP neurons caused a greater reduction of activity for the RF_{blank} condition than for the $RF_{grating}$ (11.5% reduction) or uniform (15.0% reduction) conditions ($P < 0.01$, LME, interaction between laser and stimulus; fig. S1C). Light of high intensity has been demonstrated to influence neural activity even in the absence of light-gated opsins, due to heating of the tissue (*36*). To rule out that our results could be explained by alternative causes, we performed a control experiment in which we shone light of the same laser power on the cortex in mice that did not express the optogenetic channel in VIP neurons. We did not observe a reduction in activity in the absence of optogenetic channels (Fig. 5D and fig. S1B for laminar compartments), confirming that our results were caused by the inhibition of VIP neurons in V1. The observation that VIP-mediated disinhibition is involved in contextual drive is in line with the hypothesis that feedback connections send information about the surrounding visual scene back to V1, causing enhanced activity of pyramidal neurons via a disinhibitory circuit. However, VIP inhibition did not completely abolish the gray disk responses so that our results are compatible with additional pathways that excite the cortical column, including the direct excitation of pyramidal neurons by feedback connections and routes involving subcortical brain regions.

Weak dependence of contextual responses on the orientation preference of V1 neurons

What information is contained in the response to the gray disk? We examined whether the gray disk response contained information about the orientation of the surrounding grating. In the most extreme form, the response to the gray disk might reflect a perceptual filling-in process (*37*–*39*). In filling-in, the features of surround image regions are also perceived at the blank location, such that the orientation tuning of the V1 response should reflect the orientation of the surround. We therefore determined the orientation preference of 289 strongly tuned (1-CircVar, >0.2; 289 cells; see Materials and Methods) pyramidal neurons using two-photon Ca^{2+} imaging (Fig. 6A for example neuron and Fig. 6B for distribution of orientation preference across the population). In accordance with

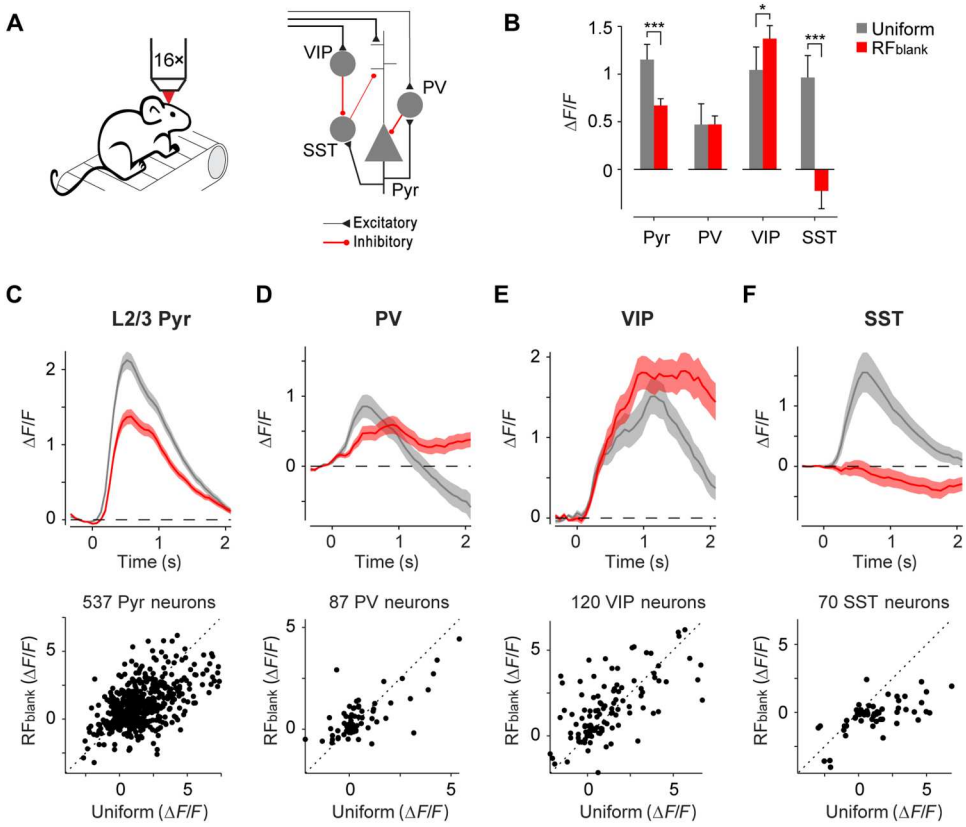

Fig. 4. Cell-specific responses to contextual stimuli in V1. (**A**) We recorded cell-specific Ca^{2+} responses of V1 neurons using two-photon microscopy (left). Schematic of V1 microcircuit involving pyramidal and the three main classes of interneurons (right). (**B**) Average activity evoked by full-screen grating (uniform, gray) or contextual (RF_{blank}, red) stimuli in a time window of 0 to 2 s in 537 pyramidal neurons (Pyr), 87 PV neurons, 120 VIP neurons, and 70 SST neurons. ***$P < 0.001$ and *$P < 0.05$, LME comparing response to uniform and RF_{blank} stimulus. (**C**) Average time course of activity across 537 L2/3 pyramidal neurons in V1 in five mice evoked by the uniform (gray) and RF_{blank} (red) stimuli (top). Activity of individual L2/3 neurons in a time window of 0 to 2 s (bottom). (**D**) Average activity of 87 PV neurons in V1 of four PV-Cre mice. (**E**) Average activity of 120 VIP neurons in V1 of four VIP-Cre mice. (**F**) Average activity of 70 SST neurons in V1 of five SST-Cre mice.

previous studies, we observed an unequal distribution of preferred orientations across the population, which was biased toward the cardinal axes (*40*, *41*). To test whether orientation-tuned neurons were more strongly driven by contextual stimuli of their preferred orientation, we normalized the response of each neuron to its preferred orientation (using full-screen gratings; see Materials and Methods) and determined the response to gratings deviating from the optimal orientation in steps of 30°. For full-screen gratings, a deviation from the preferred orientation led to a rapid drop in responses, and activity was weakest for the orthogonal orientation (Fig. 6C). We then examined the responses of the cells in the RF_{blank} condition to the different orientations in the surround of the RF. First, 66 of 289 cells had significant (bootstrap test of 1-CircVar; $P < 0.05$) orientation tuning to the grating presented in the surround of the RF compared to 156 of 289 cells that were significantly orientation tuned in the uniform condition. To examine the reliability of orientation tuning for the RF_{blank} stimulus, we fit a general linear model to each cell's responses on individual trials with surround orientation as predictor. Surround orientation explained only 5.7% of the single-trial variance (median R^2 value, 0.057; interquartile range, 0.05), whereas it explained 15% of the variance in the uniform condition (interquartile range, 12.7%).

We next examined whether neurons tuned in the uniform condition were more likely to be tuned to orientation in the RF_{blank} condition. Only 39 of the 289 cells were tuned in both conditions, a proportion that was not significantly greater than that expected by chance (H_0, 36 cells; binomial test, $P = 0.3$).

To examine the relation between orientation tuning in the two conditions, we sorted the 289 neurons based on their orientation tuning in the uniform condition, setting the preferred orientation to 0°. We then examined the tuning of the population for the RF_{blank} stimulus (Fig. 6C). There was a small but significant difference in the activity elicited by the different surround orientations (LME, $P = 0.033$; Fig. 6C), which was strongest at 0°, the preferred orientation when the RF was stimulated.

Hence, the contextual responses contain orientation information about the RF surround, but the orientation signal is considerably weaker than that when the stimulus falls in the RF. In support of this finding, we were able to decode the orientation of the surrounding grating on RF_{blank} trials at above chance performance (linear discriminant analysis, accuracy = 47.2%, chance = 16.7%, $P < 0.01$; Fig. 6D) using 80% of RF_{blank} responses to train the classifier and the 20% remaining trials for testing. The pattern of misclassifications also supported orientation tuning to the RF_{blank} stimulus because misclassifications occurred primarily at

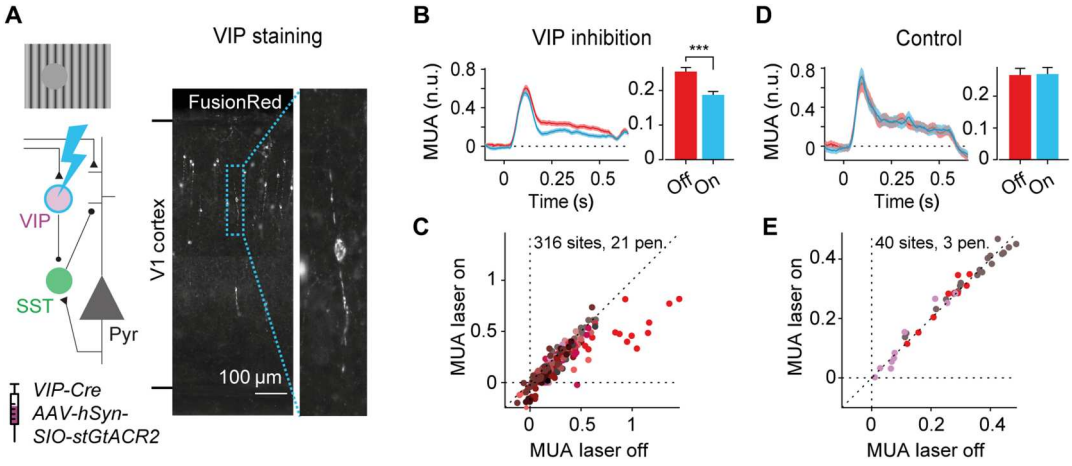

Fig. 5. Involvement of disinhibitory circuit in the contextual responses in V1. (**A**) We optogenetically inhibited VIP neurons in eight VIP-Cre mice using cell-specific expression of the inhibitory opsin stGtACR2, while electrophysiologically recording V1 activity. Left: Schematic of the disinhibitory circuit involving VIP and SST neurons. Right: Image of coronal section of V1 spanning from the bottom to the top of the cortex. Expression of the optogenetic construct is visualized by the reporter molecule FusionRed. (**B**) Average, normalized multi-unit activity (MUA) across 316 V1 recording sites evoked by the RF_{blank} stimulus, with (blue) and without (red) inhibition of VIP neurons (left). Average MUA was reduced when VIP neurons were inhibited (***$P < 0.001$, LME; time window, 0 to 0.5 s) (right). (**C**) Activity of V1 recording sites from 0 to 0.5 s after presentation of the RF_{blank} stimulus with (*y* axis) and without (*x* axis) inhibition of VIP neurons. Different colors represent different electrode penetrations. (**D** and **E**) Data from control experiment without the optogenetic construct in VIP neurons and identical laser intensities. Laser light did not affect the response to the RF_{blank} stimulus.

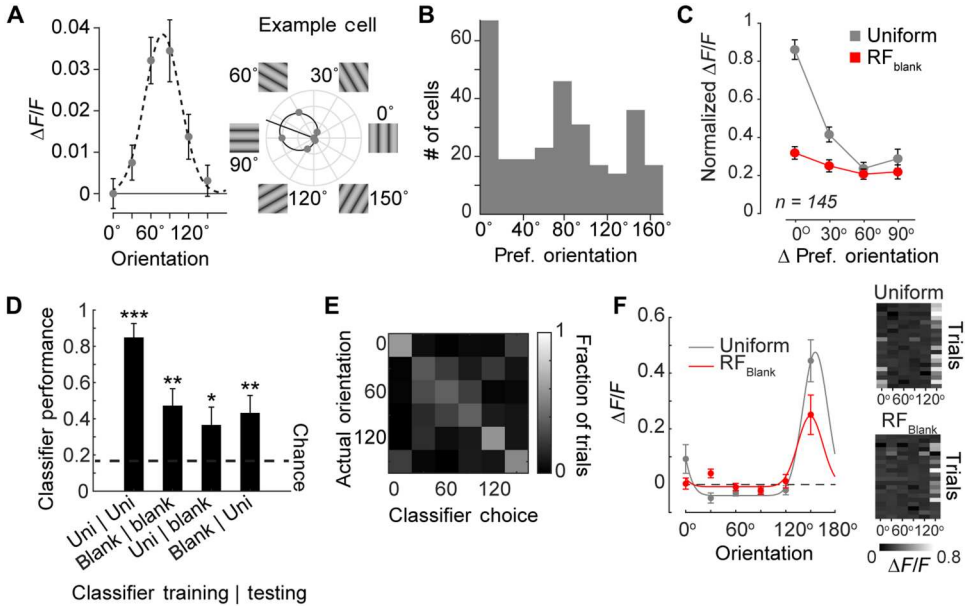

Fig. 6. Contextual response does not depend on orientation preference of V1 neurons. (**A**) Orientation tuning of an example L2/3 pyramidal neuron. Responses were evoked by static gratings with orientations ranging from 0° to 150° in steps of 30°. The dashed line shows the best-fitting circular Gaussian; the location of the peak of the Gaussian was taken as an estimate of the preferred orientation of the cell. (**B**) Distribution of preferred orientations across 289 orientation-tuned L2/3 pyramidal neurons in V1. (**C**) Average normalized response of 145 well-driven, orientation-tuned L2/3 neurons to their preferred orientation (0°), offset (30° and 60°), and orthogonal (90°) orientation to uniform grating stimuli in their RF (gray) or RF_{blank} stimuli in the surround (red). (**D**) Performance of an LDA classifier used to decode orientations. The classifier was trained/tested on either the uniform (Uni) or the RF_{blank} (blank) stimulus. We evaluated all combinations. Chance performance (dashed line) was 16.7%. The decoder performed above chance in all conditions. ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$. (**E**) Confusion matrix for the cross-validated decoder trained and tested on separate sets of RF_{blank} trials. (**F**) Response of the most informative cell for the decoder shown in (E) for the uniform and RF_{blank} stimuli. Right: Calcium activity elicited by the six different orientations (*x* axis) over 20 trials (*y* axis). The neuron's tuning for orientation was significant in the RF_{blank} condition (1-CircVar = 0.50, $P < 0.001$, bootstrap test) and the uniform condition (1-CircVar = 0.72, $P < 0.001$, bootstrap test).

neighboring orientations (Fig. 6E). The correspondence between the tuning for full-screen gratings and the RF_{blank} conditions (Fig. 6C) suggests that cross-classification might also be possible. Decoders trained on the RF_{blank} condition were able to decode orientation in the uniform condition (bootstrap test, $P < 0.001$) and vice versa (bootstrap test, $P < 0.05$), indicating that the decoder was picking up on a generalized form of orientation tuning that can also be driven from outside the RF (linear discriminant analysis, accuracy = 43.6%, $P < 0.01$). When we examined the tuning curves of cells that were most informative for the classifier's decision (see example in Fig. 6F), we observed that orientation tuning with and without RF stimulation was similar. We conclude that contextually driven responses contain a weak orientation signal, which presumably plays a role in orientation-dependent surround interactions. However, the weak orientation tuning to the RF_{blank} stimulus casts some doubt on explanations that attribute the contextual responses to perceptual filling-in. Our observations rather suggest that the gray patch survives as an independent visual element, which is not entirely filled in with the features of the surround.

Contextual responses depend on figure-ground organization

We next examined the possibility that the gray disk response is related to figure-ground organization. Could the activity elicited by the gray disk be related to the perception of a gray figure on a striped background? As a first approach, we examined whether neurons that signal a difference between a figure and the background in the RF are also the ones that are driven by the gray disk. To this aim, we presented oriented gratings with a circular figure on an orthogonal background (Fig. 7A). We used the laminar probes in V1 and ensured that the RFs were well centered and fully contained within the figure. We balanced the number of trials with the two orientations as part of the figure and background in the RF to examine figure-ground modulation, unconfounded by the orientation preference of the recorded neurons. In line with previous reports (*11*, *13*, *14*, *42*), V1 activity was enhanced for figures, compared to the background (Fig. 7B). We carried out a correlation analysis to examine whether figure-ground modulation predicts the gray disk response. The correlation coefficient, r, between the strength of figure-ground modulation and the response evoked by the RF_{blank} stimulus was 0.62 ($P < 0.001$; Fig. 7C). The correlation was strongest in the superficial layers ($r = 0.90$, $P < 0.001$), where figure-ground modulation is strong, and was weaker, but still significant, in layer 4 ($r = 0.69$, $P < 0.001$) and in the deep layers ($r = 0.31$, $P < 0.001$), where figure-ground modulation is weaker (fig. S2 shows correlations and modulation strength per laminar compartment). These results indicate a relationship between the gray disk response and figure-ground perception that was most pronounced in the superficial layers.

To further examine the relation between figure-ground perception and the gray disk response, we created stimuli for which the gray disk either was a figure occluding a background texture or part of a gray background seen through a circular cutout in a texture plane. We used dynamic occlusion cues induced by relative motion between figure and background (Fig. 7D and movie S1). We surrounded the disk with black and white dots, which provided cues about the depth relationship. In the dynamic figure condition, the dots and disk appeared in front because they were always fully visible and the relative motion caused occlusion of small parts of the texture, in a ring around the gray disk. In the dynamic ground condition, the texture acted as occluder, causing the dots to sometimes partially disappear from sight (Fig. 7D).

We electrophysiologically recorded multi-unit activity of V1 neurons in four mice across all cortical layers using laminar silicon probes. As in the previous experiments, we carefully mapped the RFs of the recorded neurons (Fig. 7E) and centered the gray disk on the RF. We applied strict inclusion criteria to ensure that the moving texture and dots never impinged upon the RF (Fig. 7E; see Materials and Methods). In line with the previous experiments and previous reports (*18*, *19*, *43*), the gray disk elicited an excitatory response at a delayed response latency (84 ± 2 ms for figures, 83 ± 2 ms for backgrounds, and 50 ± 1 ms for uniform texture stimuli; $P < 0.001$, LME). If the gray disk appeared to be a figure, then it consistently evoked more V1 activity than if it appeared as background ($P < 0.001$, LME; Fig. 7, F and G). The influence of dynamic occlusion was significant in 63 of 85 units ($P < 0.05$, t test). Compared to the peak response elicited by the full-screen stimulus (used for normalization), the strength of the effect was moderate: The average strength of the response modulation was 0.046 normalized units, and figure-ground assignment explained 4.3% of individual trial variability at the unit level. In relative terms, however, the influence of dynamic occlusion was larger because the overall response levels elicited by the dynamic figure-ground stimulus were weak. The dynamic figures produced a 48% increase in spiking activity compared to the ground condition. When dividing the units by cortical compartments, figures evoked higher responses than backgrounds in the superficial and deep layers, which receive strong feedback from higher visual areas ($P < 0.001$ for superficial and deep; Fig. 7H). In layer 4, which primarily receives feedforward input from the thalamus, the difference in response between the two stimuli was smaller and not significant ($P > 0.05$ for layer 4, LME; Fig. 7H). We note, however, that we had fewer recording sites in layer 4 than in the superficial and deep layers. Together, these results demonstrate that the activity of V1 neurons elicited by gray disk stimuli depends on figure-ground organization.

DISCUSSION

Our results provide previously unidentified insights into the mechanisms for contextual drive in V1, the activation of neurons by stimuli that do not impinge on the classical RF (*16*, *18*, *19*, *43*). We studied the profile of contextual drive across the layers of V1, the involvement of inhibitory interneurons, and the orientation tuning of the contextual response. Last, we established a relation between contextual drive and perceptual organization, the neuronal processes that segment the image into objects and the background.

In most previous studies, researchers have characterized the tuning of neurons in the visual cortex to stimuli in the classical RF, i.e., the computations enabled by the cascade of feedforward connections upstream of the neuron under study (*3*). In V1 neurons of the mouse, onset latencies are ~40 ms if the RF is stimulated (*44*). The latencies are shortest in layer 4, the input layer of the cortex, and longer in the superficial and deep layers. When we presented the gray disk stimuli, the latencies were delayed to ~80 ms, which suggests the involvement of recurrent processing routes via feedback and horizontal connections. This view was supported by the lack of the prominent current sink in layer 4 in the

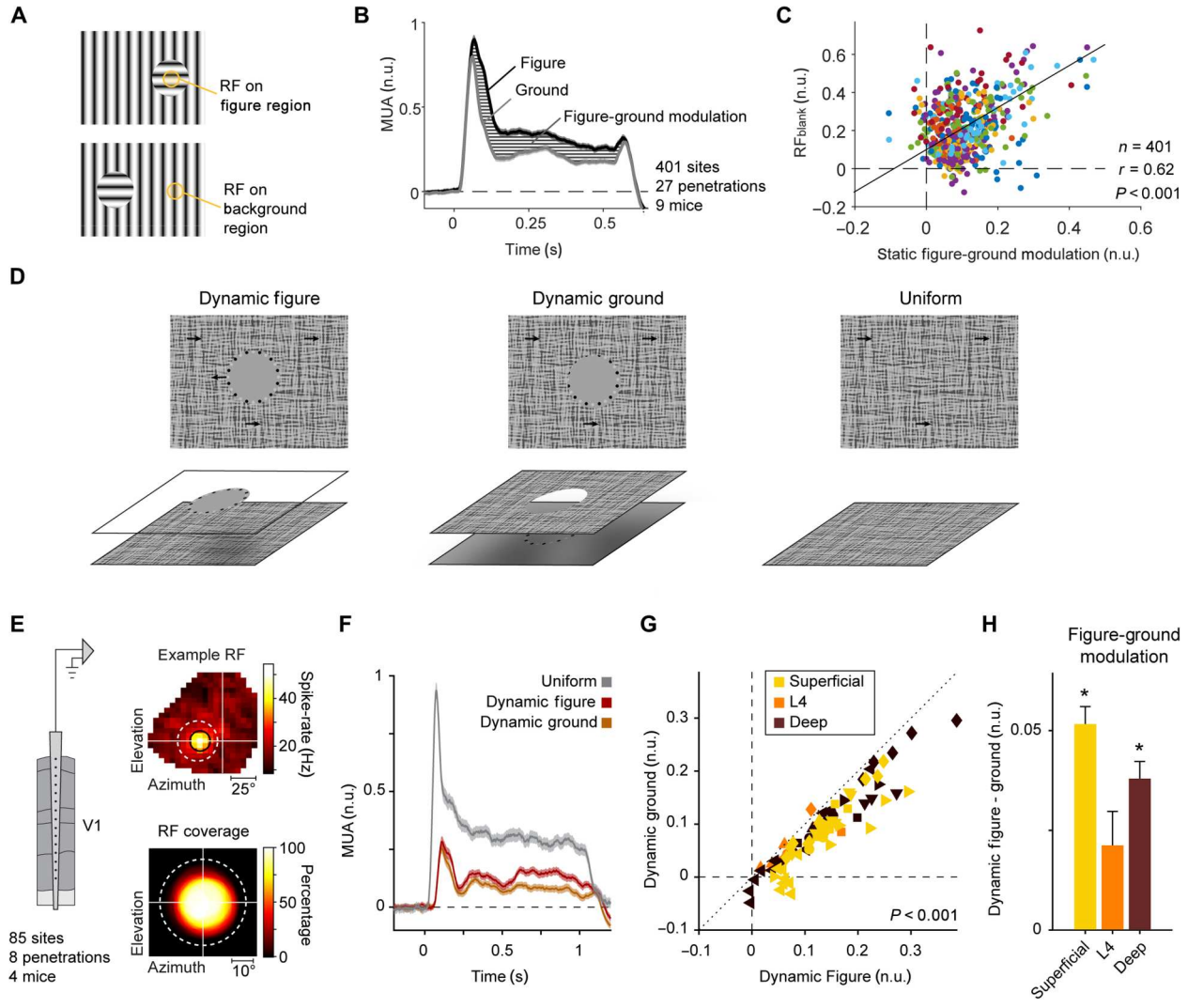

Fig. 7. Contextual responses depend on perceptual organization. (**A**) Stimuli that induce figure-ground perception consisting of circular gratings that differ from the background in orientation. We also presented stimuli for which the local orientation was orthogonal so that the orientation of the stimulus in the RF was balanced across the figure and ground conditions. (**B**) Average MUA elicited by figures and backgrounds across 401 recording sites in 27 penetrations using laminar silicon probes in V1. (**C**) Figure-ground modulation predicted the response to the RF_{blank} stimulus (401 sites, 27 penetrations, 9 mice; $r = 0.62$, $P < 0.001$, LME); different colors represent different penetrations. (**D**) Visual stimuli presented to four awake mice. We used dynamic occlusion so that the gray disk could be perceived as figure or background. (**E**) We recorded multi-unit activity in area V1 using silicon probes. Top: Example RF of one recording site. Bottom: RF coverage in percent of a population of 85 included sites relative to the figure center. (**F**) Normalized V1 activity of 85 recording sites evoked by a gray figure (red), background (orange), and a uniform full-screen texture stimulus (gray). Stimuli were presented for 1 s. (**G**) Responses evoked by the figure (*x* axis) were stronger than those evoked by the ground (*y* axis; $P < 0.001$, LMM, 85 sites; time window, 0 to 1 s); color coded by cortical compartment; different symbols represent different penetrations. (**H**) Difference in activity between the dynamic figure and dynamic ground stimuli (figure minus ground in n.u.). Values were positive on average, indicating that neural responses to figures were stronger than those to backgrounds. The effects were significant in the superficial and deep layers ($P < 0.001$, LME, 41 superficial sites, 34 deep sites; $P > 0.05$, LME, 10 sites in L4).

CSD profile of the response to the gray disk, characterizing feedforward input from the LGN. The laminar profile of activity elicited by the gray disk was "V" shaped, with the weakest activation of layer 4. This V-shaped pattern is typical for processes relying on feedback from higher areas such as attention and figure-ground modulation (*11*, *45*, *46*). Previous studies have shown that inhibiting higher visual areas leads to decreased responses to contextual stimuli (*14*, *19*), demonstrating the importance of feedback connections. Our results provide further evidence on the important role that recurrent connections play during visual processing.

We characterized the responses of the three main interneuron classes, PV, SST, and VIP neurons, to contextual stimuli. These interneurons are crucial for gating signal flow and sculpting network dynamics (*47*), and the classes differ in morphology, functional properties, connectivity, and the subcellular domains of pyramidal neurons targeted by their axons (*48*, *49*). The RF_{blank} stimulus differentially activated PV, VIP, and SST neurons. Most notably, VIP neurons were strongly driven by contextual stimuli, and SST neurons were not activated at all. This observation suggests an involvement of a previously reported canonical, disinhibitory circuit

(*23*–*27*). This disinhibitory circuit requires the activation of VIP neurons, which receive input from feedback connections (*24*) and specialize in disinhibition of pyramidal neurons through inhibition of other inhibitory neurons, in particular SST neurons (*24*, *27*, *50*, *51*). Silencing of VIP neurons decreased, but did not abolish, the contextual drive. Hence, the disinhibitory circuit contributes to activity elicited by the gray disk, but our results are compatible with additional routes including the direct excitation of pyramidal neurons by feedback connections from higher areas and input from subcortical brain regions.

In the present study, we examined whether and how the contextual drive in V1 depends on perception. One possibility was that contextual drive is related to perceptual filling-in, where blank regions in the image are filled in with features of the surround (*37*–*39*). We observed an influence of the orientation of the surround that most strongly activated V1 neurons with a matching orientation tuning. However, the small magnitude of this effect suggested that filling-in was far from complete so that the disk could survive as an independent visual object. The presence of orientation-tuned signals within the blank region indicates that feedback conveys feature-tuned information back to the primary visual cortex. Such signals could serve a variety of purposes such as computing feature differences between the RF center and surround (*2*, *19*, *29*), forming predictions about the likely RF content (*52*), and acting as selection signals to enhance the neural representation of particular features (*3*, *53*).

Another hypothesis is that the contextual drive is related to perceptual organization, the neuronal processes that group features into coherent objects and segregate them from other objects and the background (*3*, *15*, *16*, *54*–*56*). According to this hypothesis, the contextual drive relates to the perception of the gray disk as a figure superimposed on a background grating. Our results support this hypothesis, for two reasons. First, we observed a correlation between the level of figure-ground modulation elicited by orientation-defined figures and the magnitude of the gray disk response. The neurons that signaled the presence of a figure in their RF by enhancing their response were also more likely to be the ones that were strongly activated by the gray disk stimulus. Second, we used dynamic occlusion cues to create conditions in which the disk appeared as a foreground object or as part of the background. Although the possibilities to test how mice perceive such displays are limited, dynamic occlusion is a universal cue for figure-ground assignment that is present in the natural habitat of mice. We found that contextual drive was strongest when the occlusion cues signaled that the gray disk was a foreground object and weaker when they signaled that it was part of the background. These results confirm the link between contextual drive and perceptual organization.

Theories have proposed that a distinction between feedforward and feedback connections is between drivers and modulators (*3*, *57*). Feedforward connections activate neurons and determine the tuning to the information in their RFs. Feedback connections are usually thought to be merely modulatory. They can increase or decrease the feedforward response, but they usually do not drive neurons, which might otherwise give rise to hallucinations (*58*). However, a well-studied counterexample to this rule is given by illusory contours, where parallel inducers (pacmen in Fig. 1A) elicit the perception of a nonexisting line. Accordingly, illusory contours activate V1 neurons of mice and monkeys, even if there is nothing in their RF (*15*, *16*). The contextual drive elicited by gray disks represents an additional counterexample. How should we interpret the contextual drive effects that have now been shown in several studies (*13*, *18*, *19*, *43*)? Keller *et al.* (*19*) recently suggested that V1 neurons have a second RF, the fbRF, which is driven by feedback connections from neurons in higher areas that have scattered RFs. They proposed a mutually antagonistic influence between the classical RF and the fbRF, such that the fbRF is maximally stimulated when the feedforward drive of the classical RF is absent, as is the case for the gray disk stimulus. Although this idea is compatible with many of our results, we also found that subtle differences in occlusion cues, which can lead to global reorganization of the figure-ground structure, affect the strength of feedback targeted to V1. It is unclear how a "hard-wired" fbRF structure could explain such an influence of figure-ground organization. Future studies can further exploit these perceptual phenomena to deepen our understanding of how interactions between visual brain areas impose structure onto the image elements that enter into a mouse's eyes.

MATERIALS AND METHODS

Experimental model and subject details

In this study, we used 31 mice (16 males and 15 females) with an age of 2 to 12 months. All experimental procedures complied with the National Institutes of Health *Guide for Care and Use of Laboratory Animals* and the protocol was approved by the ethical committee of the Royal Netherlands Academy of Arts and Sciences and the CCD (Centrale Commissie Dierproeven). The experiments were performed in accordance with the relevant guidelines and regulations.

Method details and quantification and statistical analysis

Visual stimuli

We created the visual stimuli with the Cogent toolbox developed by the Cogent 2000 team at the Functional Imaging Labs and the ICN and Cogent Graphics developed by J. Romaya at the Laboratory of Neurobiology at the Wellcome Department of Imaging Neuroscience. We used full-screen 100% contrast sinusoidal gratings with a spatial frequency of 0.075 cycles/degree and a mean luminance of 20 cd/m^2. The textured background was generated by filtering Gaussian distributed random noise patterns through two orthogonally oriented filters. After filtering, the texture was matched to the root mean square contrast of the sine wave gratings using an iterative hard-clipping procedure with 10 iterations (*14*). We applied a previously described correction for the larger distance between the screen and the mouse at higher eccentricities (*59*). To create the contextual stimulus, we placed a 40° sized gray disk at the RF location of the recorded units. To create texture-defined figure-ground stimuli (Fig. 7A), we placed a 40° sized texture with an orthogonal orientation to the background texture at the RF location of the recorded units. In experiments in which we used dynamic occlusion to create the perception of the gray disk being part of a figure or background, the gray disk was surrounded by 24 small (1.25° radius) black and white dots. For each session, we generated new textures and used four random motion paths for the figure and background. The movement was restricted, such that the center of the disk could maximally diverge from the starting point by 3° horizontally and vertically. During electrophysiological experiments, visual stimuli were projected onto a back-projection screen placed in front of the mouse, using a DLP (Digital Light Processing)

projector (PLUS U2-X1130; size, 76 cm by 56 cm; field of view, 136° by 101.6°; resolution, 1024 × 768; and refresh rate, 60 Hz). All static visual stimuli were presented for 0.5 s with a 1.8-s ITI. Visual stimuli involving dynamic occlusion (Fig. 7, D to H) were presented for 1 s with a 1.8-s ITI. During two-photon imaging experiments, we presented visual stimuli using a 24-inch LCD monitor (resolution, 1920 × 1080; refresh rate, 60 Hz; Dell U2414H) placed in front of the left eye at an angle of 30° at a distance of 12 cm. The stimulus duration was 0.5 s with a 2.5-s ITI. We linearized the luminance profile of the screen and projector.

Surgical preparation

Anesthesia was induced in an induction box using 3 to 5% isoflurane and maintained using 1 to 2% isoflurane in an oxygen-enriched air mixture (50% air and 50% O_2). We subcutaneously injected meloxicam (5 mg/kg; 0.5 mg/ml) as general analgesic. We placed the mice in a stereotactic frame and monitored anesthesia depth by frequently checking breathing rate and paw reflexes. A feedback-controlled heating pad kept the temperature of the animal at 37°C. To prevent dehydration, we covered the eyes with ointment. The area of incision was shaved and cleaned with Betadine, and lidocaine spray was applied to the skin as a local analgesic.

Electrophysiology

Laminar silicon probe recordings. We performed acute, laminar silicon probe recordings in nine mice (five males and four females) aged 2 to 12 months. The mice were implanted with a head fixation bar and grounding pin, and we recorded neuronal activity in V1 of freely running, awake mice using laminar silicon probes (A1x32-5mm-25-177, NeuroNexus, 32 channels, 25-μm spacing). We adjusted the depth of the recording electrode with reference to the CSD (*28*). The electrical signal from the electrodes was amplified and sampled at 24.4 kHz using a Tucker-Davis Technologies recording system. To calculate the multi-unit activity (MUA), we removed muscle artifacts by re-referencing each channel to the average of all other channels before filtering the signal between 500 and 5000 Hz. We detected spikes by thresholding (positive and negative threshold) the band-passed signal at four times an estimate of the median absolute deviation and convolved the detected spikes with a Gaussian with an SD of 1.3 ms (and an integral of 1) to derive an estimate of multi-unit spike rate. We mapped the RFs using sparse noise stimuli (four 8°-by-8° checks on a grid: −64° to 16° horizontally and −22° to 66° vertically). We fit a two-dimensional (2D) Gaussian to estimate the width and center of the RF and assessed the quality of the fit using r^2 and a bootstrapped variability index (BVI). We resampled an equal number of trials (with replacement) and regenerated Gaussian fits. We calculated the BVI as the ratio of the SD of the RF center position and the SD of the fitted Gaussian. We applied the following inclusion criteria: (i) The RF completely overlapped with the figure. (ii) The RF was reliable ($r^2 > 0.33$ and BVI < 1.5). (iii) The SNR was greater than 1 [ratio of the activity evoked by a full-screen stimulus (0 to 100 ms) to the SD of baseline activity in a time window (−200 to 0 ms) relative to stimulus onset across trials]. (iv) The maximum response was greater than 2 Hz. This led to the inclusion of 401 recording sites from 27 penetrations in nine mice. To generate population responses, we subtracted the baseline activity of each recording site (−200 to 0 ms) and normalized to the peak response evoked by the full-screen texture stimulus. To assess significance, we modeled the hierarchical correlation structure in the data with LME models (*60*) because every electrode contained multiple contacts, which cannot be considered independent. We used the fitlme.m function in MATLAB for statistical testing. To test whether the RF_{blank} stimulus evoked a significant response, we fit an LME model with penetrations as a random effect and the overall intercept (the MUA to the RF_{blank} stimulus) as a fixed effect. When comparing the response to the RF_{blank} stimulus across laminar compartments, we fit an LME model with penetrations as a random effect and overall intercept and laminar compartment as fixed effects. We assessed the significance of differences across the laminar compartments with post hoc *t* tests. We estimated the response latencies by fitting a curve to the response of each recording site as described in (*34*). In short, we fitted the sum of an exponentially modulated Gaussian and a cumulative Gaussian to the time course of the response of each site. To assign the recording sites to laminar compartments, we determined a CSD profile for each recording session and located the boundary between layers 4 and 5 (*28*): superficial (100 to 400 μm above L4/5 border), layer 4 (0 to 100 μm above L4/5 border), and deep (0 to 600 μm below L4/5 border).

Silicon probe recordings paired with optogenetic inhibition of VIP neurons. We performed acute silicon probe recordings with optogenetic inhibition of VIP neurons in eight VIP-Cre mice (four males and four females) aged 2 to 12 months. We implanted the mice with a head bar and grounding screw and targeted virus injections to left V1 (2.7 mm lateral from the midline and 0.5 mm anterior of lambda) using a Nanoject III programmable injector (Drummond Scientific). We injected a total of 90 nl of AAV1-hSyn-SIO-stGtACR2-FusionRed (titer, 1×10^{13} genome copies (GC)/ml, across three depths: 200, 400, and 600 μm from dura) 3 weeks before the experiment. The virus was a gift from O. Yizhar (Addgene viral prep no. 105677-AAV1). We performed electrophysiological experiments as described above and directed an optic fiber [200 μm thickness, fiber-optic connector/physical contact (FC/PC) to 1.25-mm Ceramic Ferrule; Thorlabs, Newton] coupled to a Diode-Pumped Solid-State (DPSS) laser (BL473T3-100FC, Shanghai Lasers & Optics Century Co.) emitting blue light (wavelength, 473 nm) at the recording site. The laser was activated on 50% of the trials with a constant power of 2.83 mW/mm^2. It was turned on 200 ms before the onset of the visual stimulus, remained on for the entire stimulus presentation, and was turned off 100 ms after stimulus offset. To assess the significance of the influence of optogenetic inhibition of VIP neurons on the response to the RF_{blank} stimulus, we fit an LME model with penetrations and recording sites as random effects and overall intercept and the presence of laser light as fixed effects. We performed a control experiment to rule out tissue heating by laser light as a cause of the observed effects. To this aim, we exposed a mouse not injected with the viral construct to the same experimental conditions and repeated this experiment three times.

Silicon probe recordings while presenting figure-ground stimuli. We performed acute silicon probe recordings in nine mice (five males and four females) aged 2 to 12 months as described in the "Laminar silicon probe recordings" section above. We generated figure-ground stimuli (Fig. 7A) by placing a 40° sized circular texture region with an orientation that was orthogonal to the background texture at the RF location of the recorded units, as described in the "Visual stimuli" section above. We normalized the activity of each unit by subtracting the baseline activity in a time window (−200 to 0 ms) relative to stimulus onset and normalizing to the peak response evoked by a full-screen grating stimulus. We

quantified the strength of figure-ground modulation by subtracting the background response from the figure response and calculating the average activity in a time window from 0 to 0.5 s after stimulus onset. We used the fitlme.m function in MATLAB for statistical testing. To test whether figure-ground modulation and the RF_{blank} response were correlated, we fit an LME model with penetrations as a random effect and the overall intercept as a fixed effect.

Silicon probe recordings during dynamic occlusion stimuli. We performed acute silicon probe recordings in four C57BL/6 mice (two males and two females) as described in the "Laminar silicon probe recordings" section above. We generated visual stimuli in which dynamic occlusion was either compatible with a gray disk moving in front of a textured background or a gray background viewed through a circular hole in a textured pattern (see the "Visual stimuli" section above). In this experiment, we applied even stricter inclusion criteria to make sure that the circular inducers did not enter the RF of the recorded units. Therefore, we only included units if (i) the RF was fully contained within the figure and included a margin for the maximal movement to exclude that activity could be driven by the inducers, (ii) the RF was reliable ($r^2 > 0.33$ and BVI < 1.5), (iii) the SNR was greater than 1, and (iv) the maximum response of was greater than 2 Hz. This led to the inclusion of 85 recording sites from eight penetrations. To assess whether figures evoked significantly higher responses than backgrounds, we fit an LME model with a fixed intercept term (MUA difference between response to figure and background) and a random-effect term for the penetration using the fitlme function in MATLAB. We estimated the response latencies by fitting a curve to the response of each recording site as described in (*34*).

Two-photon imaging

Recording Ca^{2+} responses in pyramidal neurons and interneurons. We recorded Ca^{2+} responses in five Thy1-GCaMP (three males and two females), four PV-Cre (one male and 3 females), four VIP-Cre (two males and two females), and five SST-Cre (three males and two females) mice aged 2 to 12 months old. The animals were anesthetized and implanted with a head-ring. After 2 weeks, we mapped the retinotopy using pRF mapping (*61*) to determine the boundaries of visual areas and injected viral constructs in the center of V1. Thy1-GCaMP mice were additionally injected with AAV1-CaMKII-GCaMP6f-WPRE-SV40 (concentration, 10^{12} GC/ml; injection speed, 20 nl/min; depths, 600, 400, and 200 μm from dura; Penn Vector Core, University of Pennsylvania, USA) to enhance the GCaMP signal. PV-Cre, VIP-Cre, and SST-Cre mice were injected with a total of 200 to 300 nl of AAV9-CAG-flex-GCaMP6f-WPRE-SV40 (concentration, 10^{12}GC/ml; injection speed, 20 nl/min; depths: 600, 400, and 200 μm from dura; Penn Vector Core, University of Pennsylvania, USA) for cell-specific imaging in V1. We implanted double-layered glass cranial windows with a diameter of 4 to 5 mm centered on V1 and fixed the outer glass to the skull using dental cement (Vivadent Tetric EvoFlow). After 2 weeks, we habituated the mice to head immobilization on a treadmill under the two-photon microscope (Neurolabware, 16× objective (0.80 numerical aperture; Nikon); 1.7× zoom; depth, 120 to 300 μm; frame rate, 15.7 Hz; resolution, 512 × 764 pixel). A Ti-Sapphire femtosecond pulsed laser (Mai Tai, Spectra Physics) was tuned to 920 nm for delivering excitation light at a power of 20 to 50 mW.

We used the CAIMAN toolbox (*62*, *63*) for preprocessing. We performed rigid motion correction for small shifts in the data due to the motion of the animal, followed by the extraction of regions of interest (ROIs) and the $\Delta F/F$. ROI components found by the CNMF algorithm (*62*) were classified into neuronal compartments and noise using a Keras pretrained convolutional neural network classifier (CAIMAN MATLAB GitHub library; https://github.com/flatironinstitute/CaImAn-MATLAB). We only included ROIs that belonged to cell bodies in the analysis. The response of each ROI was baseline-corrected by subtracting the average $\Delta F/F$ value from the five frames before stimulus onset.

We mapped the RFs using sparse noise stimuli (12° by 12° checks on a grid: −18° to 78° horizontally, −21° to 51° vertically, 20 repetitions). Each square was flashed twice for 166 ms with a blank interval of 166 ms, followed by a delay of 500 ms. We fit a linear regression model to estimate the responses to each square, regressing out the influence of running and the interaction between the visual stimulus and running. We fit a circular 2D Gaussian to the β values to estimate the RF center and its full width at half maximum response strength. We included cells if (i) the RF was reliable ($r^2 > 0.33$, BVI < 1.5; see the "Laminar silicon probe recordings" section), (ii) the RF was well centered on the figure (Euclidean distance from RF center to figure center, <25°), and (iii) the RF completely overlapped with the figure (overlap = 100%). For VIP interneurons, we used a 50% overlap criteria because they have large RFs. We presented visual stimuli for 0.5 s with an ITI of 2.5 s to allow enough time for decay of the calcium signal of the previous trial.

For statistical analysis, we took the average baseline-corrected $\Delta F/F$ response in a window from 0.3 to 1.5 s after stimulus onset for each cell. We removed outliers using an iterative multivariate outlier removal process. The Mahalanobis distance of each cell to the mean of the full distribution across all cells was calculated and z-scored. We removed cells with an absolute z score greater than 3. If the maximum value of the pre-removal z score was greater than 6 (indicating the presence of an extreme outlier, which may distort the calculation of the z score), then the process was repeated. This procedure removed less than 3% of cells. The data from the different interneuron subclasses were concatenated together with the data from the excitatory cells in V1 before applying the outlier removal algorithm. The significance of the differences between the RF_{blank} stimulus and the full-screen grating were assessed using LMEs. The cell identity and the imaging session were included as random intercepts in the model.

Orientation tuning. In the pyramidal neuron dataset, we examined the orientation tuning of V1 neurons. To examine the orientation tuning of contextual signals, we first examined the response of pyramidal cells to full-screen gratings (the uniform condition). We used six orientations (0°, 30°, 60°, ..., 0.150°) and took the mean response of each cell between 0.3 and 1.5 s after stimulus onset as the response to each grating. We used half of the trials to measure the preferred orientation and orientation tuning strength and then examined orientation tuning responses in the held-back trials. The mean responses to the six orientations were fit with a circular Gaussian using nonlinear least-squares fitting in MATLAB as follows:

$$R(\theta) = C + R_p \cdot \exp\left(-\frac{\text{ang}_{ori}(\theta - \theta_{pref})^2}{2\sigma^2}\right)$$

where $R(\theta)$ is the response to the grating of orientation θ, C is an offset term, R_p is the response to the preferred orientation, $\text{ang}_{ori}(x) = \min[\text{abs}(x), \text{abs}(x \times 180), \text{abs}(x + 180)]$ wraps angular

differences to the interval 0° to 90°, and σ is the SD of the Gaussian. Tuning strength was assessed using 1-CircVar, which is a robust estimator of tuning strength (*64*). It was calculated as follows

$$1 - \text{CircVar} = \left| \frac{\sum_k R(\theta_k) e^{(2i\theta_k)}}{\sum_k \text{abs}[R(\theta_k)]} \right|$$

where $R(\theta_k)$ is the response to the orientation θ_k (in radians). The significance of the tuning strength was assessed by bootstrapping. For each cell, we randomly resampled with replacement an equal number of trials as in the original dataset but shuffled the responses across orientations and calculated 1-CircVar 10,000 times to create the expected null distribution of the 1-CircVar statistic from which we derived the *P* value. To calculate the orientation tuning curve of the RF_{blank} condition, we recoded orientation as the difference between the orientation shown and the preferred orientation of the cell (measured using half the trials from the uniform condition). We rounded the preferred orientation to the nearest 30° before calculating the difference to yield four values (0, 30, 60, and orthogonal). The data were normalized to the mean response of the population in the uniform condition at the preferred orientation (using half the trials). We discarded cells that were weakly tuned (1-CircVar, <0.3) or had a weak or negative response to the uniform condition at their preferred orientation [max response, <0.028 (median response)], leaving 146 cells for this analysis.

To decode the orientation of the surrounding grating using the contextual response, we used linear discriminant analysis (LDA) using the function fitcdiscr.m in MATLAB. We used all pyramidal cells that passed our general inclusion criteria (see above; $n = 578$). We trained each LDA decoder by randomly selecting 80% of the trials and then testing the decoder on the remaining 20% of trials. No training of hyperparameters was performed, and we used the default values of the function. We built two classifiers, one trained on the uniform condition and one trained on the RF_{blank} condition. Each classifier was then tested on both conditions to yield a total of four classifier accuracy scores. To estimate the variability of the decoders, we repeated the procedure 1000 times with random selection of trials. The most informative cell (Fig. 6F) was taken as the one with the highest linear discriminant contrast (*65*).

Histology

To examine virus expression, we deeply anesthetized the mice with Nembutal and transcardially perfused them with phosphate-buffered saline (PBS), followed by 4% paraformaldehyde (PFA) in PBS. We extracted the brain and postfixed it overnight in 4% PFA before moving it to a PBS solution. We cut the brains into 50-μm-thick coronal slices and mounted them on glass slides. We imaged the slices on a Zeiss Axioplan 2 microscope (20× objective, Zeiss Plan-Apochromat, 0.63 NA) using custom-written Image-Pro Plus software.

Supplementary Materials

This PDF file includes:
Figs. S1 and S2

Other Supplementary Material for this manuscript includes the following:
Movie S1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. J. Allman, F. Miezin, E. McGuinness, Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* **14**, 105–126 (1985).
2. J. J. Knierim, D. C. van Essen, Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* **67**, 961–980 (1992).
3. P. R. Roelfsema, Cortical algorithms for perceptual grouping. *Annu. Rev. Neurosci.* **29**, 203–227 (2006).
4. H. Adesnik, W. Bruns, H. Taniguchi, Z. J. Huang, M. Scanziani, A neural circuit for spatial summation in visual cortex. *Nature* **490**, 226–231 (2012).
5. C. D. Gilbert, T. N. Wiesel, Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133 (1983).
6. W. Bair, J. R. Cavanaugh, J. A. Movshon, Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J. Neurosci.* **23**, 7690–7701 (2003).
7. L. Nurminen, S. Merlin, M. Bijanzadeh, F. Federer, A. Angelucci, Top-down feedback controls spatial summation and response amplitude in primate visual cortex. *Nat. Commun.* **9**, 2281 (2018).
8. J. Vangeneugden, E. H. van Beest, M. X. Cohen, J. A. M. Lorteije, S. Mukherjee, L. Kirchberger, J. S. Montijn, P. Thamizharasu, D. Camillo, C. N. Levelt, P. R. Roelfsema, M. W. Self, J. A. Heimel, Activity in lateral visual areas contributes to surround suppression in awake mouse V1. *Curr. Biol.* **29**, 4268–4275.e7 (2019).
9. J. J. Nassi, S. G. Lomber, R. T. Born, Corticocortical feedback contributes to surround suppression in V1 of the alert primate. *J. Neurosci.* **33**, 8504–8517 (2013).
10. V. A. F. Lamme, P. R. R. Roelfsema, The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* **23**, 571–579 (2000).
11. V. A. Lamme, The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15**, 1605–1615 (1995).
12. M. W. Self, J. K. Possel, D. Jeurissen, P. R. Roelfsema, J. C. Peters, J. Reithler, R. Goebel, P. Ris, L. Reddy, S. Claus, J. C. Baayen, The effects of context and attention on spiking activity in human early visual cortex. *PLOS Biol.* **14**, e1002420 (2016).
13. U. H. Schnabel, C. Bossens, J. A. M. Lorteije, M. W. Self, H. Op de Beeck, P. R. Roelfsema, Figure-ground perception in the awake mouse and neuronal activity elicited by figure-ground stimuli in primary visual cortex. *Sci. Rep.* **8**, 17800 (2018).
14. L. Kirchberger, S. Mukherjee, U. H. Schnabel, E. H. van Beest, A. Barsegyan, C. N. Levelt, J. A. Heimel, J. A. M. Lorteije, C. van der Togt, M. W. Self, P. R. Roelfsema, The essential role of recurrent processing for figure-ground perception in mice. *Sci. Adv.* **7**, eabe1833 (2021).
15. T. S. Lee, M. Nguyen, Dynamics of subjective contour formation in the early visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1907–1911 (2001).
16. A. Pak, E. Ryu, C. Li, A. A. Chubykin, Top-down feedback controls the cortical representation of illusory contours in mouse primary visual cortex. *J. Neurosci.* **40**, 648–660 (2020).
17. P. Kok, F. P. De Lange, Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Curr. Biol.* **24**, 1531–1535 (2014).
18. T. S. Lee, D. Mumford, R. Romero, V. A. F. Lamme, The role of the primary visual cortex in higher level vision. *Vision Res.* **38**, 2429–2454 (1998).
19. A. J. Keller, M. M. Roth, M. Scanziani, Feedback generates a second receptive field in neurons of the visual cortex. *Nature* **582**, 545–549 (2020).
20. R. P. N. Rao, D. H. Ballard, Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
21. K. Friston, A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* **360**, 815–836 (2005).
22. A. Derrington, Vision: Filling in and pop out. *Curr. Biol.* **6**, 141–143 (1996).
23. M. M. Karnani, J. Jackson, I. Ayzenshtat, A. H. Sichani, K. Manoocheri, S. Kim, R. Yuste, Opening holes in the blanket of inhibition: Localized lateral disinhibition by VIP interneurons. *J. Neurosci.* **36**, 3471–3480 (2016).
24. S. Zhang, M. Xu, T. Kamigaki, J. P. P. Hoang Do, W.-C. Chang, S. Jenvay, K. Miyamichi, L. Luo, Y. Dan, Long-range and local circuits for top-down modulation of visual cortex processing. *Science* **345**, 660–665 (2014).
25. S. Lee, I. Kruglikov, Z. J. Huang, G. Fishell, B. Rudy, A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nat. Neurosci.* **16**, 1662–1670 (2013).
26. A. Guet-McCreight, F. K. Skinner, L. Topolnik, Common principles in functional organization of VIP/calretinin cell-driven disinhibitory circuits across cortical areas. *Front. Neural Circuits* **14**, 1–14 (2020).
27. C. K. Pfeffer, M. Xue, M. He, Z. J. Huang, M. Scanziani, Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* **16**, 1068–1076 (2013).
28. U. Mitzdorf, Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. *Physiol. Rev.* **65**, 37–100 (1985).

29. M. W. Self, J. A. M. Lorteije, J. Vangeneugden, E. H. van Beest, M. E. Grigore, C. N. Levelt, J. A. Heimel, P. R. Roelfsema, Orientation-tuned surround suppression in mouse visual cortex. *J. Neurosci.* **34**, 9290–9304 (2014).
30. K. S. Rockland, D. N. Pandya, Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* **179**, 3–20 (1979).
31. D. J. Felleman, D. C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
32. G. Blasdel, J. Lund, Termination of afferent axons in macaque striate cortex. *J. Neurosci.* **3**, 1389–1413 (1983).
33. V. K. Berezovskii, J. J. Nassi, R. T. Born, Segregation of feedforward and feedback projections in mouse visual cortex. *J. Comp. Neurol.* **519**, 3672–3683 (2011).
34. J. Poort, F. Raudies, A. Wannig, V. A. F. Lamme, H. Neumann, P. R. Roelfsema, The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* **75**, 143–156 (2012).
35. H. Dana, T.-W. Chen, A. Hu, B. C. Shields, C. Guo, L. L. Looger, D. S. Kim, K. Svoboda, Thy1-GCaMP6 transgenic mice for neuronal population imaging in vivo. *PLOS ONE* **9**, e108697 (2014).
36. S. F. Owen, M. H. Liu, A. C. Kreitzer, Thermal constraints on in vivo optogenetic manipulations. *Nat. Neurosci.* **22**, 1061–1065 (2019).
37. H. Komatsu, The neural mechanisms of perceptual filling-in. *Nat. Rev. Neurosci.* **7**, 220–231 (2006).
38. P. De Weerd, R. Gattass, R. Desimone, L. G. Ungerleider, Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature* **377**, 731–734 (1995).
39. M. Meng, D. A. Remus, F. Tong, Filling-in of visual phantoms in the human brain. *Nat. Neurosci.* **8**, 1248–1254 (2005).
40. A. K. Kreile, T. Bonhoeffer, M. Hübener, Altered visual experience induces instructive changes of orientation preference in mouse visual cortex. *J. Neurosci.* **31**, 13911–13920 (2011).
41. M. M. Roth, F. Helmchen, B. M. Kampa, Distinct functional properties of primary and posteromedial visual area of mouse neocortex. *J. Neurosci.* **32**, 9716–9726 (2012).
42. M. W. Self, D. Jeurissen, A. F. van Ham, B. van Vugt, J. Poort, P. R. Roelfsema, The segmentation of proto-objects in the monkey primary visual cortex. *Curr. Biol.* **29**, 1019–1029.e4 (2019).
43. W. Li, P. Thier, C. Wehrhahn, Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *J. Neurophysiol.* **83**, 941–954 (2000).
44. R. Land, G. Engler, A. Kral, A. K. Engel, Response properties of local field potentials and multiunit activity in the mouse visual cortex. *Neuroscience* **254**, 141–151 (2013).
45. M. W. M. W. Self, T. van Kerkoerle, H. Supèr, P. R. P. R. Roelfsema, Distinct roles of the cortical layers of area V1 in figure-ground segregation. *Curr. Biol.* **23**, 2121–2129 (2013).
46. T. van Kerkoerle, M. W. Self, P. R. Roelfsema, Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nat. Commun.* **8**, 13804 (2017).
47. R. Tremblay, S. Lee, B. Rudy, GABAergic Interneurons in the neocortex: From cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
48. X. Jiang, S. Shen, C. R. Cadwell, P. Berens, F. Sinz, A. S. Ecker, S. Patel, A. S. Tolias, Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).
49. Y. Kubota, Untangling GABAergic wiring in the cortical microcircuit. *Curr. Opin. Neurobiol.* **26**, 7–14 (2014).
50. F. Hajós, K. Zilles, A. Schleicher, M. Kálmán, Types and spatial distribution of vasoactive intestinal polypeptide (VIP)-containing synapses in the rat visual cortex. *Anat. Embryol.* **178**, 207–217 (1988).
51. H. J. Pi, B. Hangya, D. Kvitsiani, J. I. Sanders, Z. J. Huang, A. Kepecs, Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521–524 (2013).
52. G. B. Keller, T. D. Mrsic-Flogel, Predictive processing: A canonical cortical computation. *Neuron* **100**, 424–435 (2018).
53. J. H. R. Maunsell, S. Treue, Feature-based attention in visual cortex. *Trends Neurosci.* **29**, 317–322 (2006).
54. D. H. Grosof, R. M. Shapley, M. J. Hawken, Macaque VI neurons can signal "illusory" contours. *Nature* **365**, 550–552 (1993).
55. F. W. Smith, L. Muckli, Nonstimulated early visual areas carry information about surrounding context. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20099–20103 (2010).
56. T. D. Albright, G. R. Stoner, Contextual influences on visual processing. *Annu. Rev. Neurosci.* **25**, 339–379 (2002).
57. S. M. Sherman, R. W. Guillery, On the actions that one nerve cell can have on another: Distinguishing "drivers" from "modulators". *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7121–7126 (1998).
58. H. Neumann, A. Yazdanbakhsh, E. Mingolla, Seeing surfaces: The brain's vision of the world. *Phys. Life Rev.* **4**, 189–222 (2007).
59. J. H. Marshel, M. E. Garrett, I. Nauhaus, E. M. Callaway, Functional specialization of seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
60. H. Scheffe, A "mixed model" for the analysis of variance. *Ann. Math. Stat.* **27**, 23–36 (1956).
61. E. van Beest, S. Mukherjee, L. Kirchberger, U. H. Schnabel, C. van der Togt, R. R. M. Teeuwen, A. Barsegyan, A. Meyer, J. Poort, P. Roelfsema, M. W. Self, Mouse visual cortex contains a region of enhanced spatial resolution. *Nat. Commun.* **12**, 4029 (2020).
62. E. A. Pnevmatikakis, D. Soudry, Y. Gao, T. A. Machado, J. Merel, D. Pfau, T. Reardon, Y. Mu, C. Lacefield, W. Yang, M. Ahrens, R. Bruno, T. M. Jessell, D. S. Peterka, R. Yuste, L. Paninski, Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* **89**, 285–299 (2016).
63. E. A. Pnevmatikakis, A. Giovannucci, NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *J. Neurosci. Methods* **291**, 83–94 (2017).
64. M. Mazurek, M. Kager, S. D. Van Hooser, Robust quantification of orientation selectivity and direction selectivity. *Front. Neural Circuits.* **8**, 1–17 (2014).
65. A. Walther, H. Nili, N. Ejaz, A. Alink, N. Kriegeskorte, J. Diedrichsen, Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage* **137**, 188–200 (2016).

Acknowledgments: We thank the mechatronics and animal departments at the Netherlands Institute for Neuroscience for technical support, E. Ruimschotel and C. Levelt for help with mouse breeding, and the members of the Roelfsema laboratory for discussions of this project. **Funding:** This project was supported by the European Union's Horizon 2020 and FP7 Research and Innovation Program (grants 720270 and 785907), grant agreement 650003 "Human Brain Project," ERC grant agreement 339490 "Cortic_al_gorithms," FLAG-ERA JTC grant "ChampMouse," and the Erasmus Mundus "NeuroTime" program. **Author contributions:** L.K., M.W.S., and P.R.R. designed the study. L.K. and M.W.S. conducted electrophysiological experiments and analyzed the data. S.M. performed the two-photon experiments, and M.W.S. performed the analysis. L.K., M.W.S., and P.R.R. wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All data and the computer code used to analyze the data can be accessed via DOI: 10.25493/6GCC-3XH. Correspondence and requests for materials can be sent to P.R.R. (p.roelfsema@nin.knaw.nl).

Submitted 2 June 2022
Accepted 19 December 2022
Published 20 January 2023
10.1126/sciadv.add2498