NEUROSCIENCE

The essential role of recurrent processing for figure-ground perception in mice

Lisa Kirchberger[1†], Sreedeep Mukherjee[1†], Ulf H. Schnabel[1†], Enny H. van Beest[1†], Areg Barsegyan[1], Christiaan N. Levelt[2,3], J. Alexander Heimel[4], Jeannette A. M. Lorteije[5], Chris van der Togt[1], Matthew W. Self[1‡], Pieter R. Roelfsema[1,6,7*‡]

The segregation of figures from the background is an important step in visual perception. In primary visual cortex, figures evoke stronger activity than backgrounds during a delayed phase of the neuronal responses, but it is unknown how this figure-ground modulation (FGM) arises and whether it is necessary for perception. Here, we show, using optogenetic silencing in mice, that the delayed V1 response phase is necessary for figure-ground segregation. Neurons in higher visual areas also exhibit FGM and optogenetic silencing of higher areas reduced FGM in V1. In V1, figures elicited higher activity of vasoactive intestinal peptide–expressing (VIP) interneurons than the background, whereas figures suppressed somatostatin-positive interneurons, resulting in an increased activation of pyramidal cells. Optogenetic silencing of VIP neurons reduced FGM in V1, indicating that disinhibitory circuits contribute to FGM. Our results provide insight into how lower and higher areas of the visual cortex interact to shape visual perception.

INTRODUCTION

Neurons at early processing levels of the visual system initially analyze the visual scene in a fragmented manner. They carry out a local analysis of the image elements in their small receptive fields (RFs) and feed the visual information forward to higher visual areas (HVAs). Neurons in HVAs have larger RFs and integrate information to represent increasingly abstract features of the visual scene, including object category and identity (*1*, *2*). However, there are many images for which the analysis is not complete when information has reached the HVAs (*3*). It has been hypothesized that these images require the recirculation of activity back to lower areas through recurrent connections, which include feedback connections from higher to lower visual areas and horizontal connections between neurons within the same cortical area. These recurrent routes are associated with additional synaptic and conduction delays, so that the recurrent influences usually are not expressed when visual cortical neurons are initially activated by the stimulus but during a later phase of their response (*4*).

One of the proposed roles for recurrent processing is that it supports perceptual organization, the process that groups image elements of behaviorally relevant objects and segregates them from other objects and the background (*5*). An advantage of recurrent processing for perceptual organization is that neurons in low-level visual cortical areas have small RFs so that they represent the segmentation results at a high spatial resolution. Figure 1A illustrates a few example figure-ground images that induce perceptual organization; a subset of the image elements forms a figure, which is segregated from the background. Studies in monkeys demonstrated that neuronal responses in the primary visual cortex (area V1) that are elicited by figures are stronger than those elicited by background regions, even if the image elements in the RF are the same (compare the RF for stimulus 1 versus 2 and 3 versus 4 in Fig. 1B) (*4*, *6*–*8*). This relative enhancement of neural activity on figures is known as figure-ground modulation (FGM). FGM has been measured with functional magnetic resonance imaging and electroencephalography in the visual cortex of human participants (*9*), and a study in a human subject documented that image elements of a figure also elicit stronger spiking activity than background elements in low-level areas of the visual cortex (*10*).

FGM is a contextual effect (it originates from outside the neurons' RFs) that occurs later than the initial response elicited by stimulus onset, suggesting that it represents a recurrent influence on the neuronal response, which originates from HVAs (*4*, *11*). In accordance with an important role of feed back connections, FGM in V1 is stronger when a monkey pays attention to a figure (*12*) or if it enters into the subject's awareness (*13*). However, direct evidence for a causal role of feedback connections in FGM has been lacking. Specifically, it is unknown whether FGM is just an epiphenomenon of processing in HVAs that feed back to V1 or if the delayed V1 response phase is necessary for perception. Hence, three of our aims are to test whether the late V1 response is necessary for figure-ground perception, whether FGM occurs in HVAs, and whether HVAs are the source of FGM in V1.

A fourth aim was to gain insight into the roles of interneurons in figure-ground segregation. A previous study proposed that the activity of somatostatin (SOM)–expressing inhibitory neurons (SOM neurons) is a candidate mechanism for background suppression (*14*). SOM neurons inhibit the dendrites of pyramidal neurons and might detect homogeneous image regions that are likely to be part of the background. Another study (*15*) suggested a mechanism by which feedback connections could increase the activity for relevant

[1]Department of Vision and Cognition, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA, Amsterdam, Netherlands. [2]Molecular Visual Plasticity Group, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA, Amsterdam, Netherlands. [3]Department of Molecular and Cellular Neuroscience, Center for Neurogenomics and Cognitive Research, VU University, Amsterdam, Netherlands. [4]Cortical Structure and Function Group, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA, Amsterdam, Netherlands. [5]Cognitive and Systems Neuroscience Group, Swammerdam Institute for Life Sciences, Faculty of Science, University of Amsterdam, 1098XH Amsterdam, Netherlands. [6]Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU University, Amsterdam, Netherlands. [7]Department of Psychiatry, Academic Medical Center, Amsterdam, Netherlands.
*Corresponding author. Email: p.roelfsema@nin.knaw.nl
†These authors contributed equally to this work.
‡These authors contributed equally to this work as joint last authors.

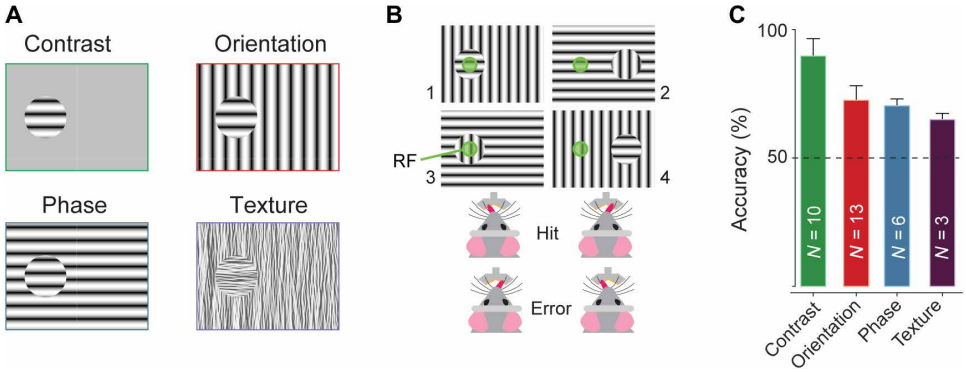

Fig. 1. Mice segregate different types of figures from a background. (**A**) Stimuli that induce figure-ground perception. The stimulus either contained a grating that differed from the background in orientation (top right) or phase (bottom left). For comparison, we presented a circular grating on a gray background ("contrast," top left) that did not require figure-ground segregation. We also presented a texture stimulus with an orientation difference between figure and ground (bottom right). (**B**) The mice reported the figure location by licking a spout on the corresponding side. The image elements falling in the RF (green circles) were identical in the figure and background conditions so that differences in neuronal activity could only be caused by the context determined by stimulus regions outside the RF. (**C**) Average accuracy of the mice during all recording sessions included in this study. Error bars, SD across mice. The number of mice per task is indicated within each bar. Accuracies of individual mice are presented in table S1.

representations in lower areas via vasoactive intestinal peptide–positive neurons (VIP neurons). VIP neurons are interneurons that inhibit SOM cells, and they thereby disinhibit the cortical column (*16*–*18*). However, the activity of interneurons elicited by figure-ground stimuli and their involvement in generating FGM in V1 remain to be demonstrated.

Previous studies in freely moving mice demonstrated that mice can perceive figures that differ from the background in orientation or phase (top right and bottom left in Fig. 1A) (*19*, *20*). To address our questions, we developed a figure-ground paradigm for head-fixed mice, to have sufficient control over the stimulus position relative to the mouse and simultaneously measure brain activity. Furthermore, we tested the generality of figure-ground perception by also including stimuli where the figure differed from the background in texture (bottom right in Fig. 1A), similar to those used in previous monkey studies (*6*).

We report that HVAs are crucial for the generation of FGM in V1 and that figure-ground perception fails if the late V1 response phase is inhibited. Furthermore, our results demonstrate that figures enhance VIP cell activity and suppress the activity of SOM cells. We report that VIP neurons cause disinhibition, which is more pronounced at the figure location so that optogenetic suppression of VIP cell activity reduces FGM. These insights into the neuronal mechanisms for figure-ground perception are likely to generalize to other cases in which the neuronal representations of simple features in lower brain areas and more abstract features in higher brain areas need to interact with each other during perception.

RESULTS

Mice can segregate figures from the background using various cues

We trained head-restrained mice to report whether a circular grating (0.075 cycles/deg) or texture pattern was positioned on the left or right side of a display by licking one of two spouts (Fig. 1, A and B). We rewarded correct choices with a drop of liquid reward. In the initial training phase, mice learned to report the location of figures on a gray background. We then gradually increased the contrast of the background grating until it equaled the contrast of the figure grating, and the only remaining difference was orientation or phase. Mice reached an accuracy of 90 ± 7% (mean ± SD; N = 10 mice) in the task with a gray background, 72 ± 6% (N = 13 mice) if the figure differed in orientation from the background, and 70 ± 3% if it differed in phase (N = 6 mice) (Fig. 1C). The accuracy of the mice was significantly above chance level (ps < 0.001 for all mice and tasks, Bonferroni-corrected one-sided binomial test; table S1). To test generalization, we presented a new figure-ground stimulus with orthogonally oriented textures (Fig. 1A, bottom right) and observed immediate generalization with an average accuracy of 65 ± 2% (N = 3 mice, ps < 0.001 for all mice). The demonstration that head-restrained mice can report figure-ground structure elicited by various cues implies that the many powerful methods available to measure and change neuronal activity in this species can be used to study the neuronal mechanisms of perceptual organization.

FGM in V1 is evoked by various visual cues

To examine V1 activity elicited by these different figure-ground stimuli, we recorded multiunit activity (MUA) with multichannel silicon probes (N = 198 recording sites in 13 electrode penetrations in six awake mice, passively viewing the stimuli). For the figure-ground stimuli, we ensured that the stimulus in the RF was identical in the figure and ground conditions (Figs. 1B and 2A), so that activity differences could only be caused by contextual information from stimulus regions outside the classical RF. For comparison, we also presented contrast stimuli (Fig. 1A, top left) where the RF either fell on a grating or the gray background. We quantified FGM for individual recording sites with the d-prime, a measure of reliability of the difference between the figure and ground response on single trials (time window, 100 to 500 ms after stimulus onset). Activity elicited by orientation, phase, and texture-defined figures was stronger than that elicited by backgrounds, despite the sensory stimulus in the RF being identical (Fig. 2B; figure S1 shows responses of example recording sites). We used linear mixed-effects models (see Materials and Methods) to account for the fact that recording sites from the same penetration typically have higher correlations than those from different penetrations (*21*) and observed significant

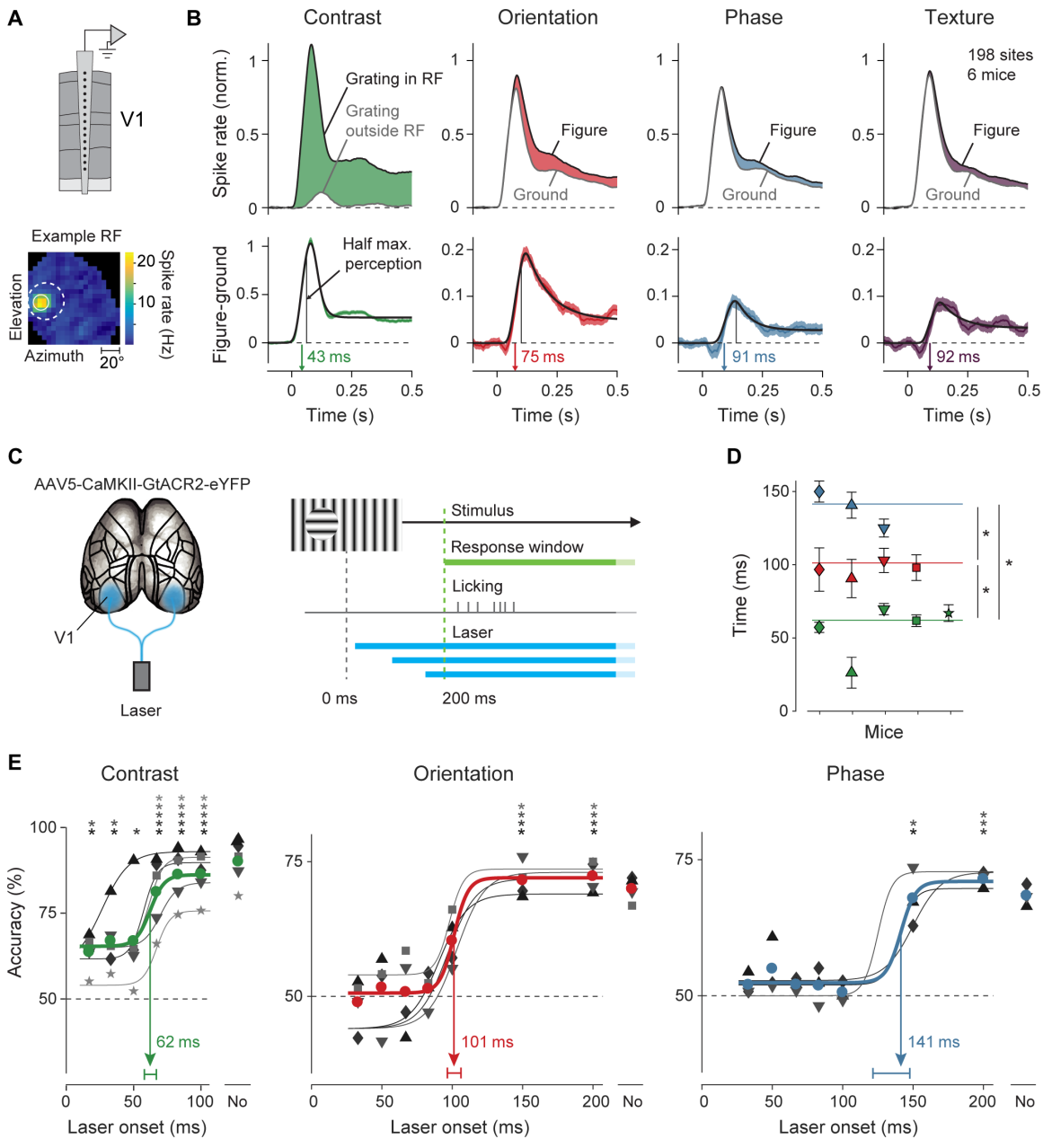

Fig. 2. Figure-ground perception relies on late V1 activity. (**A**) Silicon probe recordings in V1 and an example RF determined with a mapping stimulus (see Materials and Methods). White continuous circle, estimated RF boundary. Dashed circle, figure location. (**B**) Average MUA response elicited by a contrast-defined figure and the figure-ground stimuli (198 recording sites in 6 mice). Bottom: The difference in activity elicited by figure and ground (FGM) (visually driven response for the contrast stimulus) and a curve that was fitted to determine FGM latency (visual response latency for the contrast stimulus; see Materials and Methods). The black vertical lines illustrate the V1 processing time required for half-maximal accuracy [from (D) and (E)]. (**C**) We optogenetically inhibited V1 activity in both hemispheres at different delays after stimulus appearance using blue laser light. The mice reported figure location after 200 ms by licking. (**D**) V1 processing time necessary to reach half-maximal performance in the contrast detection task (green) and for orientation-defined (red) and phase-defined (blue) figure-ground stimuli. Symbols show data from individual mice, and the horizontal lines show the average latencies. Error bars, SEM determined by bootstrapping. (**E**) Accuracy of five mice in the tasks, as function of the onset time of optogenetic V1 silencing. The data points on the right of the graphs are the accuracies on trials without silencing. The data of individual mice are shown as gray/black symbols [same symbols per mouse as in (D)] and curves are fits of logistic functions. Colored curves, fits to the average accuracy across mice. The inflection point of the curve was taken as measure of latency (the 95% confidence interval is shown above the abscissa). Asterisks above or to the right of the plots indicate when the accuracy of the individual mice was above chance level (*$P < 0.05$, Bonferroni-corrected binomial test).

FGM for the three types of figure-ground stimuli (all ps < 0.001). The average d-prime was 1.65 for the contrast stimulus (indexing the reliability of the visual response), and it was 0.44, 0.22, and 0.24 for the orientation-defined, phase-defined, and texture stimuli, respectively (indexing FGM). FGM was stronger for orientation-defined figures than for figures with a phase or texture difference (ps < 0.001, Bonferroni-corrected post hoc *t* tests). To determine FGM latency, we subtracted V1 activity elicited by the background from that elicited by the figure and fitted a curve to the response difference (bottom in Fig. 2B; see Materials and Methods). We measured a visual response latency of 43 ms with the contrast-defined stimulus. FGM latencies for all figure-ground stimuli were delayed relative to the visual response latency and only occurred during a later processing phase (all ps < 0.001, bootstrap test). The latency of FGM for orientation defined figure-ground stimuli was 75 ms. Latencies for the phase-defined (91 ms) and texture figure-ground stimuli (92 ms) were, in turn, longer than for the orientation-defined figure-ground stimulus (both ps ≤ 0.001, bootstrap test). We also measured FGM in V1 of four mice performing the figure detection task (fig. S2). The strength of FGM on correct trials was similar to that during passive viewing and even to that on error trials.

Late V1 activity is necessary for figure-ground segregation

FGM in V1 occurs during the late, sustained processing phase. To determine whether late V1 activity is necessary for figure-ground perception, we inhibited V1 neurons at different latencies while mice were performing the figure-ground segregation and contrast detection task. We expressed the inhibitory opsin GtACR2, a light-activated chloride channel (*22*), bilaterally in V1 pyramidal neurons in five mice. Blue light quickly and reliably inhibited the neuronal activity across all cortical layers (fig. S3, G to I). We determined how long V1 activity was essential for reliable performance, by varying the onset latency of optogenetic inhibition across trials (Fig. 2C) (*23*). Once V1 activity was inhibited at a given latency, the blue light remained on until the end of the trial. In the contrast detection task, silencing from 33 ms after stimulus onset onward, which abolished the entire V1 response, reduced the accuracy of the mice (Fig. 2E) (Bonferroni-corrected one-sided binomial test, $P < 0.05$ in each mouse). In most animals, accuracy remained above chance level, which suggests that a relatively low accuracy in the contrast detection task can be maintained by brain regions outside V1, although we cannot exclude the possibility that V1 silencing was incomplete in some of the mice. The accuracy quickly recovered when we postponed optogenetic silencing until after the visually driven V1 response, reaching its half-maximal value after 62 ms (95% confidence interval, 58 to 67 ms) (Fig. 2, D and E). In the orientation-defined figure-ground task, silencing of V1 reduced the accuracy to chance level, and it only recovered to its half-maximal value when V1 silencing was postponed to 101 ms (95% confidence interval, 97 to 106 ms). The required processing time was even longer (141 ms) if the figure was defined by a phase difference (95% confidence interval, 121 to 148 ms; ps < 0.05 for latency differences between all conditions; Fig. 2D). We also measured the minimal reaction time (mRT) for the different stimuli, as the 10-ms bin of the reaction time distribution at which the mice made more correct than erroneous licks (*24*). For contrast-defined figures, the mRT was 245 ms, and it was 255 and 315 ms for the orientation- and phase-defined figure-ground stimuli, respectively (fig. S3, A to C). We found that mRTs for phase-defined figures were longer than for orientation-defined figures (paired *t* test, $P < 0.05$; fig. S3B), in accordance with the longer minimally necessary V1 processing time for these stimuli.

We ensured that the blue light itself did not cause interference by carefully shielding the mouse's eyes from the blue light that shone on the cortex and placing a second blue light-emitting diode (LED) below the mouse's head that flashed at random intervals, which the animal learned to ignore (see Materials and Methods). Furthermore, shining blue light on cortical areas not expressing the optogenetic channel did not reduce the accuracy in the task (fig. S3, D and E).

The duration of V1 activity required for accurate task performance might also depend on task difficulty (*23*) because the accuracy of the mice on the figure-ground tasks was lower than that on the contrast task. We carried out two control analyses to examine the influence of task difficulty on our measure of the minimally necessary processing time. First, FGM occurred later for figure-ground displays in which the figure differed from the background in phase than when it differed in orientation (Fig. 2B), whereas task difficulty for these stimuli (i.e., accuracy of the mice) was similar (Figs. 1C and 2, D and E). The minimally necessary processing time for phase-defined figures (141 ms) was markedly longer than that for orientation-defined figures (101 ms) ($P < 0.05$, bootstrap test), a difference that cannot be explained by task difficulty effects. Second, two of the mice attained a relatively low accuracy in the contrast task, which was comparable to the accuracy of the other mice in the figure-ground tasks (fig. S3F). Nevertheless, the minimal V1 processing time was short in these two mice, and they reached their relatively low ceiling performance if we started V1 inhibition after only 70 to 90 ms. These results indicate that it is unlikely that variations in difficulty were responsible for the differences between minimal V1 processing times.

We conclude that the initial V1 response transient suffices for accurate performance in the contrast detection task, whereas the late V1 response phase is necessary for figure-ground segregation, supporting the hypothesis that it is read out by brain structures that select the appropriate motor response (*25*). The latencies derived from the optogenetic experiment were systematically longer than the neuronal latencies (compare neuronal FGM latencies to time of half-maximal accuracy; black lines in the bottom panel of Fig. 2B), suggesting that downstream areas need to accumulate visually driven activity in the contrast task and to accumulate FGM during a longer epoch for figure-ground displays to support accurate performance.

FGM is present in multiple higher cortical areas

The late onset of FGM in V1 suggests that it may be a result of recurrent interactions between V1 and HVAs. To examine neuronal activity across the cortical areas that could provide feedback to V1 (*26*), we used wide-field imaging in nine Thy1-GCaMP6f 5.17 mice with a transparent skull (*27*, *28*). We first measured the retinotopy to determine the boundaries between the areas and to align them with the Allen Brain Map (see Materials and Methods and figs. S4 and S5C) and investigated FGM elicited by orientation-defined ($N = 9$ mice) and phase-defined ($N = 5$ mice) figure-ground displays, during passive viewing. We used the measured retinotopy to transform the V1 activity profile into visual field coordinates, and we observed that the entire visual field region occupied by the figure was labeled with enhanced activity, with a protracted time course caused by the dynamics of the calcium sensor (time window, 150 to 300 ms; paired *t* test, $P < 0.001$) (Fig. 3, A and B, and fig. S5, A and B). Both orientation- and phase-defined figures elicited more activity

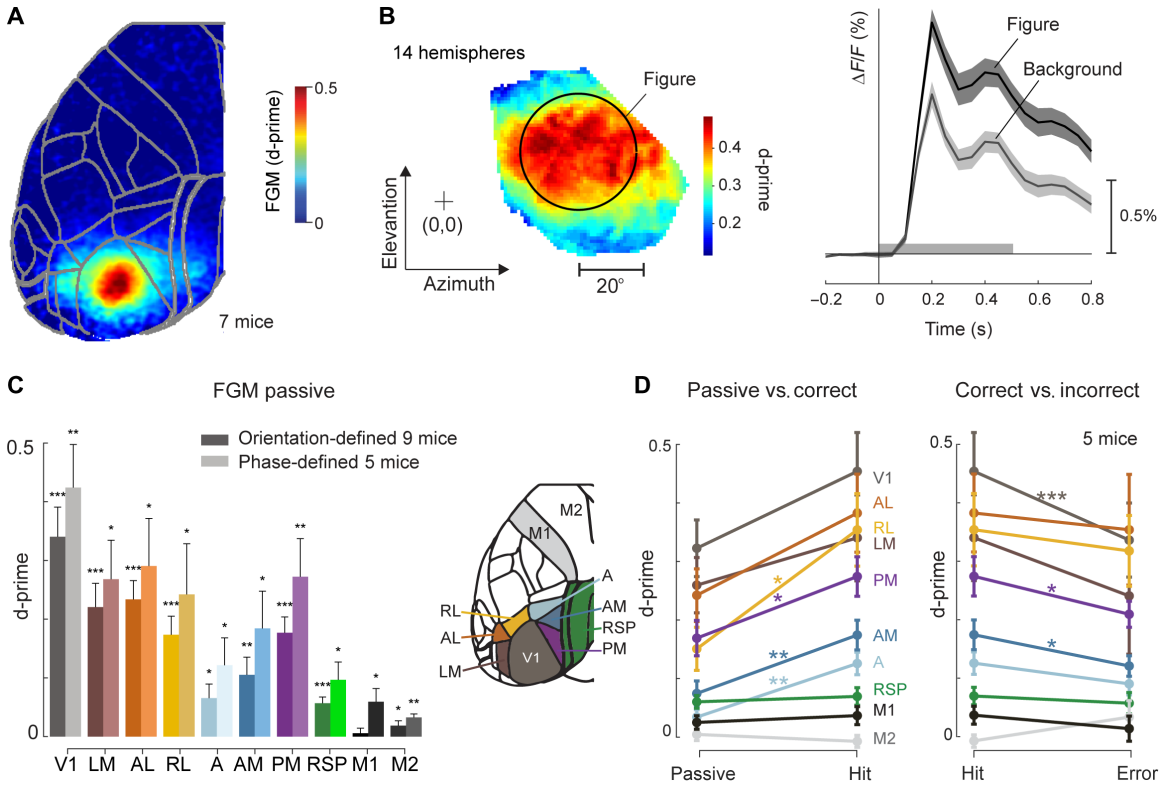

Fig. 3. FGM across areas of the visual cortex measured with wide-field imaging. (**A**) Strength of FGM for orientation-defined figures quantified using the d-prime during passive viewing, averaged across all hemispheres at 150 to 300 ms after stimulus onset. We used seven mice, two hemispheres per mouse (two mice were not used in this analysis because we used a different figure position in these animals). Data of individual mice is presented in fig. S5A. (**B**) Left: The average V1 d-prime values over 14 hemispheres from seven mice elicited by an orientation defined figure-ground stimulus transformed into visual coordinates, based on the retinotopy. Data from the right visual field was reflected horizontally and projected into the left visual field before averaging. The black circle illustrates the outline of the figure appearing on the screen in front of the mouse. Plus symbol, center of the visual field. The colored region indicates the extent of the visual field, which could be reliably mapped in at least five mice per hemisphere. Right: Average time course of the response of V1 pixels falling in the figure representation (black circle in the left) when presented with an orientation-defined figure (black curve) or background (gray curve) (for the response phase-defined stimuli, see fig. S5B). The shaded area denotes SEM across hemispheres. (**C**) We aligned the brains of individual mice to the Allen Brain common coordinate framework (see Materials and Methods) and computed d-prime (averaged across all pixels within each area) for orientation and phase-defined figure-ground stimuli in passively viewing mice. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, post hoc t tests. Error bars denote SEM. (**D**) D-prime for orientation defined figure-ground stimuli during passive viewing, hits and errors in the figure detection task. Error bars denote SEM ($N = 5$ mice). *$P < 0.05$, **$P < 0.01$, ***$P < 0.005$.

than the background, and the level of FGM differed between areas (orientation: $F_{9,72} = 27.1$, $P < 0.001$; phase: $F_{9,36} = 16.7$, $P < 0.001$; Fig. 3C). There were no significant differences between the levels of FGM elicited by orientation- and phase-defined stimuli across the five mice that were trained on both types of stimuli (fig. S5D).

To examine whether FGM increases during active perception, we trained five mice to perform the orientation figure-ground task during wide-field imaging. FGM during correct figure detection trials was generally stronger than during passive viewing [two-way analysis of variance (ANOVA) with factors area and task; interaction, $F_{18,36} = 1.961$, $P < 0.05$; Fig. 3D] and specifically for areas A, AM, PM, and RL, in accordance with a previous study demonstrating that activity in some of these areas increases when mice engage in a visual task (*29*). We also compared the activity between correct and erroneous trials. Just as in our electrophysiological experiments, FGM was present when mice made an error. In some areas (AM, PM, and V1) FGM was stronger for correct responses than during errors (Fig. 3D). These results, together, indicate that FGM is a robust phenomenon that occurs in different behavioral states.

Wide-field imaging pools calcium signals across multiple cellular compartments, with possible contributions of axons that originate in other brain regions, and the method emphasizes layer 1 activity (*30*). We, therefore, also examined the calcium signals of cell bodies of excitatory neurons within area V1 in five mice using two-photon imaging (Fig. 4A). The mice were injected with the virus AAV1-CaMKII-GCaMP6f-WPRE-SV40 and passively viewed stimuli with figures defined by an orientation difference. Just as in the electrophysiological (Fig. 2B) and wide-field recordings (Fig. 3), V1 responses elicited by figures were elevated relative to those elicited by the background (Fig. 4, C to E; linear mixed-effects model, see Materials and Methods; 1122 cells, $P < 0.001$). We could use a longer time window (300 to 1500 ms) than in the wide field experiments, because the mice passively viewed the stimuli.

We recorded calcium signals of cell bodies in HVAs: LM, AL, RL, AM, and PM (Fig. 4B). FGM was particularly strong in LM (Fig. 4F), its strength in AL was comparable to that of V1, and it was weaker in RL. These results confirm that neurons in multiple HVAs exhibit FGM (linear mixed-effects models in five mice; LM: 429

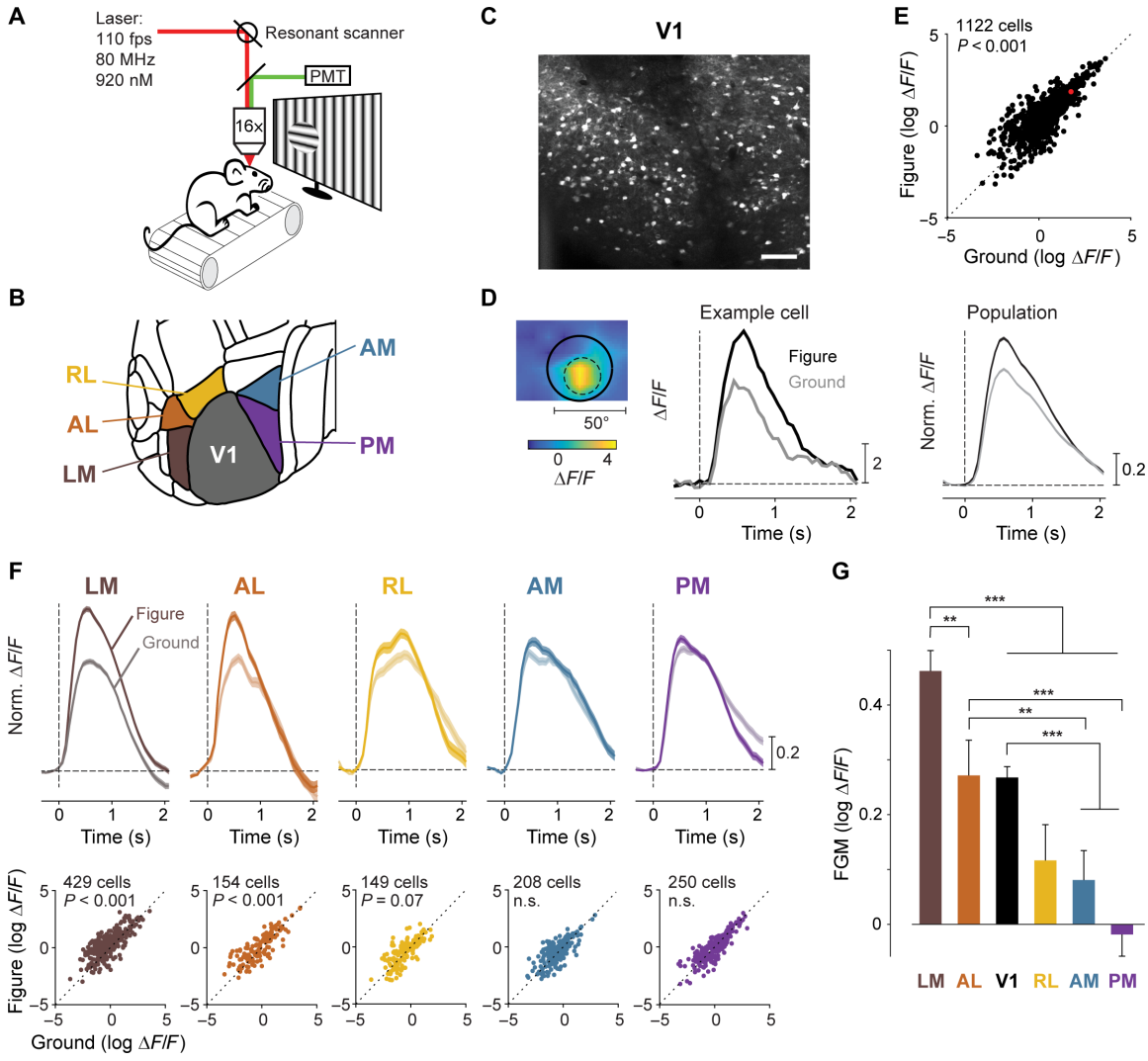

Fig. 4. Ca^{2+} response of pyramidal cells in HVAs during figure-ground segregation. (**A**) We used two-photon imaging to measure the Ca^{2+} responses of pyramidal neurons in five Thy1-GCaMP6f mice passively viewing orientation-defined figure-ground stimuli. (**B**) Organization of visual areas in the mouse cortex. (**C**) Example correlation image; luminance values represent the average correlation between a pixel and its eight neighbors across 3000 frames. Scale bar, 100 μm. (**D**) Left: Example RF (dashed circle) and figure position (continuous circle). Middle: Ca^{2+} activity of an example pyramidal neuron. Right: Average activity across the population of 1122 pyramidal cells in V1 elicited by the figure-ground stimuli, after normalization to the maximum of the figure condition. The shaded area shows ±SEM but is difficult to see given its small magnitude. (**E**) The activity of pyramidal cells in V1 in a time window from 300 to 1500 ms after stimulus onset. Because the data are positively skewed, we plot log (Δ*F*/*F*). Figures elicited stronger responses than the background. Red symbol, example cell of (D); note that the log transform causes the symbol to be just above the diagonal, despite substantial FGM. (**F**) Ca^{2+} responses of cell bodies of excitatory cells in higher areas LM, AL, RL, AM, and PM, in the same mice as (E). Top: Normalized population activity of excitatory cells evoked by figures (saturated colors) and backgrounds (less saturated colors) in normalized units. Shaded areas, SEM. Bottom: Log (Δ*F*/*F*) of individual excitatory cells (time window, 300 to 1500 ms after stimulus onset) evoked by figures (*y* axis) and the background (*x* axis). n.s., not significant. (**G**) Comparison of FGM in six visual areas based on log-transformed activity. FGM was strongest in LM, followed by AL and V1, weaker in RL, and weakest in AM and PM (Bonferroni-corrected post hoc tests; $^{**}P < 0.01$ and $^{***}P < 0.001$).

cells, $F_{1,856} = 154$, $P < 0.001$; AL: 154 cells, $F_{1,306} = 17.9$, $P < 0.001$; RL: 149 cells, $F_{1,296} = 3.2$, $P = 0.07$). FGM in AM and PM neurons was even weaker during passive viewing (both ps > 0.1; see Fig. 4G for comparisons between areas). We note, however, that the wide-field imaging results revealed AM and PM as areas in which FGM increases during task performance (Fig. 3D). Furthermore, the comparison of two-photon and wide-field imaging results supports the view that wide-field imaging emphasizes layer 1 activity (*30*), because LM cell bodies had stronger FGM than V1 cell bodies and they provide strong input into layer 1 of area V1 (*31*). We also investigated whether RF size predicted the strength of FGM, but we did not observe such a relationship (fig. S5E).

HVAs contribute strongly to FGM in V1

Figures cause stronger activity than backgrounds in HVAs. We next investigated whether feedback from HVAs could be a source of the increased V1 activity that is elicited by figures, during the delayed response phase. Alternative possibilities are that FGM is generated locally, within V1 (*32*) or that it reflects an influence from subcortical visual areas, mediated via the thalamus. To examine the contribution

of cortical feedback to FGM in V1, we silenced HVAs (Fig. 5A). We injected AAV1-CaMKII-stGtACR2-FusionRed medially (in areas PM, AM, and A) and laterally from V1 (LM, RL, and AL) in Thy1-GCaMP6f mice (*27*). The virus encodes a soma-targeted version of the inhibitory opsin GtACR2 (*33*), which was expressed in excitatory neurons. We ensured that the virus did not spread into V1 by identifying the area borders with population RF mapping and targeted the viral injections at a minimal distance of 0.5 mm from the V1 border (Fig. 5, B and C). We used laminar silicon probes to record V1 activity elicited by figure-ground stimuli and optogenetically inhibited neuronal activity in the HVAs on 50% of the trials, from 200 ms before stimulus onset until 100 ms after stimulus offset. Optogenetic inhibition did not have a systematic influence on the early peak response of V1 neurons (time window, 0 to 100 ms; fig. S6A), but it reduced the late V1 responses elicited by figures and backgrounds (fig. S6C). This V1 activity reduction was more pronounced for figures than for backgrounds, decreasing FGM by 55, 100, and 79% for the orientation defined, phase-defined, and textured figure-ground stimuli, respectively (Fig. 5, D to L, and fig. S6D) (five mice, eight penetrations, 109 recording sites, time window of 100 to 500 ms, linear mixed-effects model, $P < 0.001$ for orientation-defined and $P < 0.05$ for phase- and texture-defined figures). Hence, feedback from HVAs enhances the representation of the figure in V1 more than the representation of the background, and it thereby accounts for a large fraction of the FGM in V1.

Roles of the different interneuron classes in V1 FGM

We next asked how the feedback signal from HVAs might influence activity in the cortical microcircuit in V1 to enhance the activity evoked by figures. Recent studies suggested that the different types of interneurons have unique roles in controlling the activity of the cortical column (*16*). We therefore examined how the three main interneuron subclasses in V1 respond to figure-ground stimuli and compared their activity to that of pyramidal cells. Feedback axons can enhance V1 activity by directly targeting excitatory neurons and suppress activity by contacting parvalbumin (PV) or SOM-expressing interneurons (Fig. 6A) (*15*, *34*). A previous study demonstrated that SOM neurons suppress V1 activity elicited by larger, homogeneous image regions (*14*), but their response to figure-ground displays has not yet been studied. Last, feedback connections to V1 could also disinhibit the cortical column by targeting VIP interneurons, which inhibit SOM cells and thereby cause disinhibition of pyramidal cells (Fig. 6A) (*15*–*18*, *35*). The relative contribution of these excitatory, inhibitory, and disinhibitory motifs to perceptual organization is unknown, and modeling studies indicate that the interactions between the cell types might be complex, given the large number of influences between the different cell types (*36*). Nevertheless, a few predictions can be made. First, VIP neurons are predicted to be more active for the figure than for the background. Second, their targets, SOM neurons should respond less to the figure than to the background, so that FGM should be inverted for these cells.

To image the activity of VIP neurons, we expressed GCaMP6f in VIP-Cre mice and observed that figures elicited stronger VIP neuron activity than backgrounds ($N = 155$ neurons in four mice; linear mixed-effects model; $F_{1,308} = 9.1$, $P < 0.01$) (Fig. 6, C to E), consistent with the hypothesis that these neurons receive excitatory feedback from HVAs about salient visual stimuli (*15*). To examine the activity of SOM cells, we expressed GCaMP6f in SOM-Cre mice and found that activity elicited by figures was weaker than that elicited by the background ($N = 178$ neurons in 5 mice, $F_{1,354} = 7.6$, $P < 0.01$) (Fig. 6, F to H). This result supports the hypothesis that the release of SOM inhibition contributes to the extra activity elicited by the figure (Fig. 6A) (*18*). We also examined the activity of PV-positive cells, which have been implicated in feedforward inhibition and control of the gain of the cortical column (*16*). We expressed GCaMP6f in V1 of PV-Cre mice and observed that figures evoked stronger responses in PV cells than backgrounds ($N = 159$ neurons in four mice, $F_{1,316} = 119$, $p < 0.001$) (Fig. 6, I to K). The level of FGM differed significantly between the different cell types (linear mixed-effects model across cell types, including the excitatory cells in Fig. 4E; interaction between figure-ground condition and cell type, $F_{1,3220} = 32.5$, $P < 0.001$). FGM was strongest for PV cells, weaker and similar in magnitude in excitatory cells and VIP interneurons, and inverted in the SOM population (see Fig. 6B for post hoc comparisons of the different cell classes). Visually evoked responses are modulated by locomotion (*37*). We observed that the activity of excitatory neurons, PV-, VIP-, and SOM-positive interneurons were all enhanced by running (linear mixed-effects model; main effect of running, all ps < 0.001), but there was no significant interaction between running and FGM in any of the cell types ($P > 0.05$ for all cell types).

VIP neurons contribute to FGM in V1

Do VIP neurons indeed enhance the representation of figures in V1 as predicted by their proposed role in silencing SOM neurons? To test the involvement of this disinhibitory circuit, we inhibited the activity of VIP neurons and measured how it influences FGM in V1. We expressed an inhibitory opsin (stGtACR2) in VIP neurons in V1 and suppressed their activity while electrophysiologically recording V1 activity evoked by orientation-defined figures and the background (Fig. 6L and fig. S7E). We started the inhibition of VIP neurons 250 ms after stimulus onset, to selectively influence the late response phase during which HVAs feedback to V1 and to prevent interference with the early phase during which V1 neurons propagate activity to HVAs. We found that inhibiting VIP-activity reduced FGM in V1, decreasing the d-prime by 47% (Fig. 6, M and N; linear mixed-effects model, $P < 0.05$, three mice, eight penetrations, 168 recording sites, time window of 250 to 500 ms). This finding supports the role of VIP neurons in generating FGM in V1.

The relationship between FGM and surround suppression

In addition to their involvement in figure-ground perception, feedback connections are thought to play a role in surround suppression, i.e., the reduction of activity if the features in the RF of V1 neurons match those in the surround (*14*, *38*, *39*). In an extreme scenario, surround suppression might fully account for FGM, strongly suppressing V1 activity evoked by the large, homogeneous background and weakly suppressing activity evoked by the smaller figure (*40*). Is FGM indeed the mere consequence of surround suppression?

The results of the experiment in which we silenced higher areas (Fig. 5) provided evidence against this scenario. Silencing of higher areas reduced V1 activity evoked by the background and activity evoked by the figure even more (fig. S6C), which is evidence against a predominantly suppressive feedback effect. At first sight, these results appear to be at odds with a previous study (*41*) demonstrating a suppressive influence of HVAs on V1 activity elicited by backgrounds. However, that study focused on neurons with strong surround

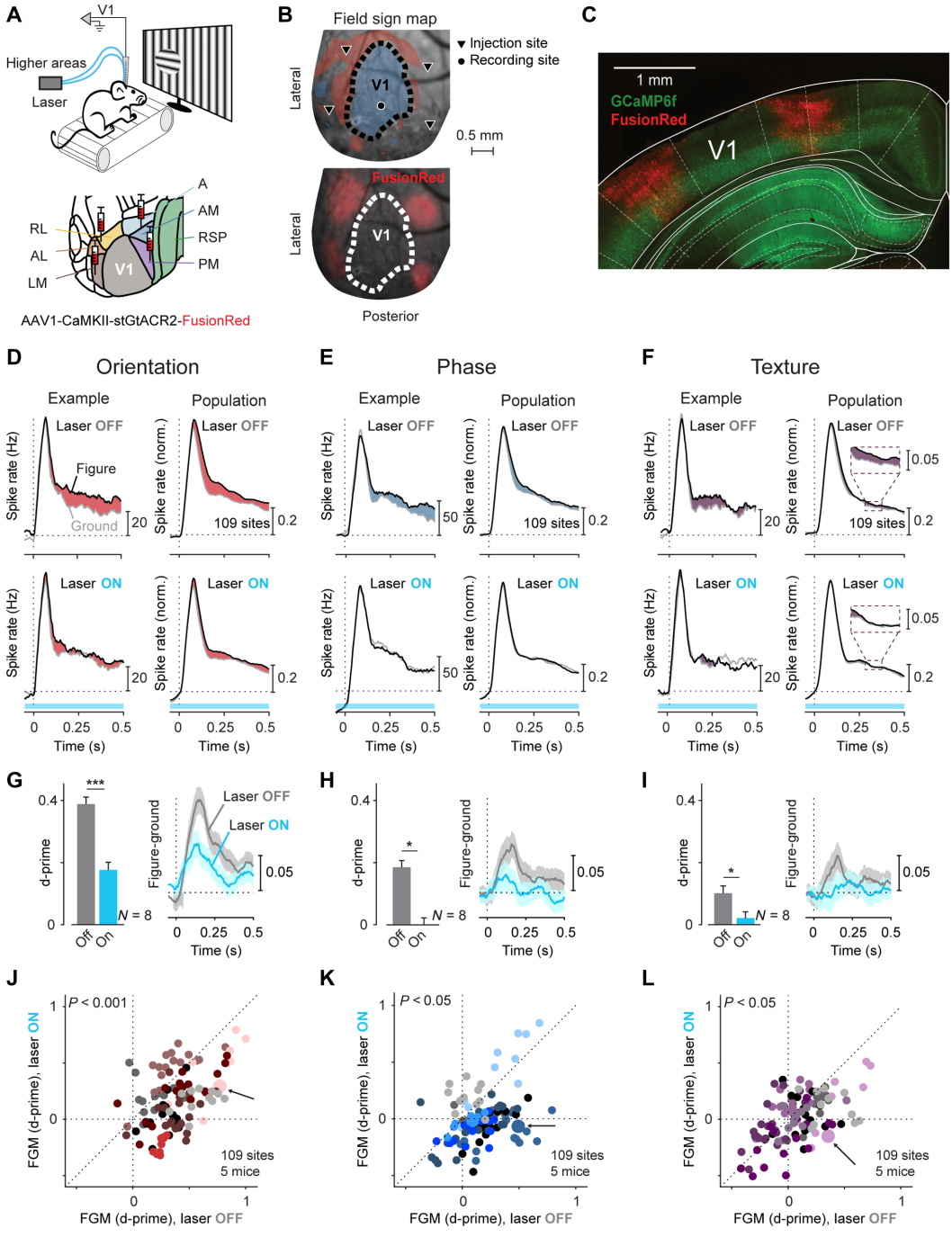

Fig. 5. Contribution of feedback from HVAs to FGM in V1. (**A**) We optogenetically inhibited HVAs using stGtACR2 while recording V1 activity with laminar electrodes. (**B**) Top: Example field sign map generated using population RF mapping overlaid on the visual cortex (see Materials and Methods). Blue (red) regions have a mirror-inverted (noninverted) retinotopy. Black circle, location of an example electrode penetration, guided by the field sign map and the blood vessel pattern. Triangles, virus injection sites, minimally 500 μm from the V1 boundary (dashed line). Bottom: Virus expression profile. (**C**) Coronal brain slice showing stGtACR2 (red) and GCaMP6f (green) expression aligned to the Paxinos and Watson adult mouse atlas. (**D**) MUA at an example V1 recording site (left) and across a population of 109 sites (eight electrode penetrations and five mice) (right) elicited by an orientation-defined figure (black curves) and the background (gray curves) without (top) and with optogenetic silencing (bottom) of HVAs. Data for population responses was normalized to the laser off condition (see Materials and Methods). The example site is marked in (J) with an arrow and a larger symbol. (**E** and **F**) MUA elicited by phase-defined (E) and texture-defined figure ground stimuli (F). (**G** to **I**) Left: FGM quantified as d-prime (time window, 100 to 500 ms after stimulus presentation) in V1 was lower during optogenetic inhibition of HVAs (linear mixed-effects model, see Materials and Methods) for orientation-defined (G), phase-defined (H), and textured figure-ground stimuli (I); $^{*}P < 0.05$, $^{***}P < 0.001$. Right: Average time course of FGM across 109 sites without (gray curve) and with inhibition of activity in HVAs (blue curve). (**J** to **L**) FGM d-prime with (*y* axis) and without (*x* axis) optogenetic inhibition of HVAs for orientation defined (J), phase-defined (K), and textured stimuli (L). Data of different penetrations are shown in distinct colors (linear mixed-effects model, $P < 0.001$ for orientation and $P < 0.05$ for phase and textured figure-ground stimuli).

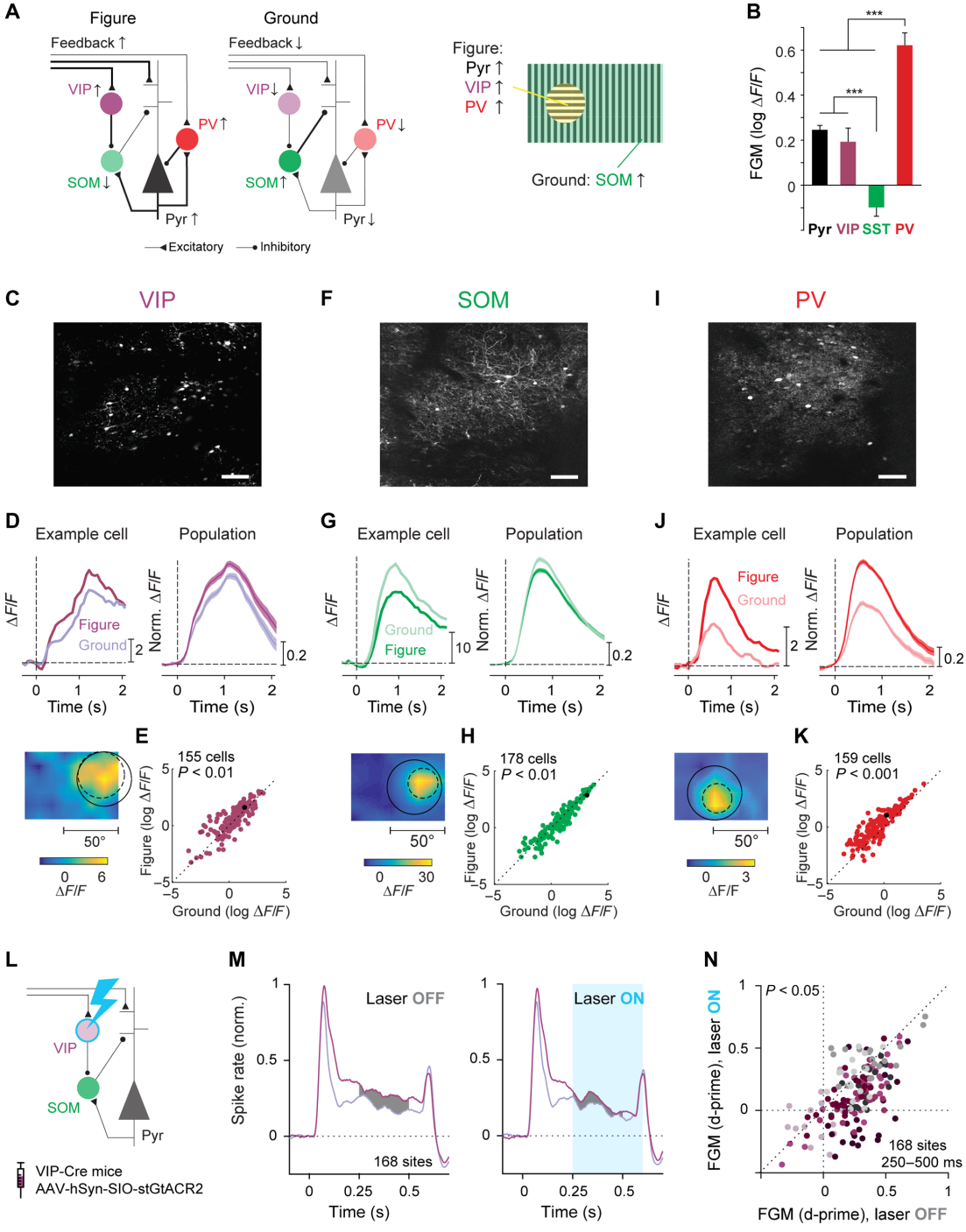

Fig. 6. The activity of interneurons in area V1 during figure-ground segregation. (**A**) Left: Schematic of the cortical microcircuit during presentations of figures and backgrounds. Right: Our results reveal that figures enhance the activity of pyramidal neurons, VIP-, and PV cells, whereas SOM neurons are more active for the background. (**B**) Significant FGM differences between neural classes (linear mixed-effects model: Bonferroni-corrected post hoc comparisons, ***$P < 0.001$). (**C**) GCaMP6f expression in an example VIP-Cre mouse induced with AAV9-CAG-Flex-GCaMP6f. (**D**) Activity of an example VIP neuron and a population of 155 VIP neurons in four mice elicited by the figure-ground stimuli. Bottom: The RF of the example cell. (**E**) Figures elicited stronger responses in VIP neurons than the background (linear mixed-effects model, $P < 0.01$). Black symbol, the example cell from (C). (**F** to **H**) Responses of SOM neurons evoked by figure-ground stimuli (178 neurons, five mice). The calcium response elicited by the background was stronger than that elicited by the figure (linear mixed-effects model, $P < 0.01$). (**I** to **K**) Responses of PV-neurons to figure-ground stimuli (159 neurons, four mice). Figures elicited stronger calcium responses than the background (linear mixed-effects model, $P < 0.001$). (**L**) We optogenetically inhibited VIP neurons in V1 using stGtACR2 while recording V1 activity with laminar electrodes. (**M**) Average, normalized activity of a population of 168 recording sites (eight penetrations, three mice) elicited by an orientation-defined figure (dark purple) and background (light purple) without (left) and with (right) inhibition of VIP neuron activity, starting 250 ms after stimulus onset. (**N**) FGM, quantified using d-prime with (*y* axis) and without (*x* axis) optogenetic inhibition of VIP neurons. Data of different penetrations are shown in distinct colors (linear mixed-effects model, $P < 0.05$; time window, 250 to 500 ms).

suppression, whereas we included all neurons here. We therefore examined whether the influence of HVA silencing on V1 activity elicited by background regions depends on the level of surround suppression. HVA silencing increased the activity elicited by the background for neurons with strong surround suppression (fig. S6, E and F), in accordance with the previous study (*41*), but it decreased it when all neurons were taken into account. In other words, feedback connections have a net excitatory influence on V1 activity elicited by the background but decrease activity of the subset of neurons with strong surround suppression.

Surround suppression can be separated into two components: untuned suppression and tuned suppression. Untuned suppression does not depend on stimulus features and therefore cannot account for the FGM because figure-ground stimuli always had the same size. Tuned suppression is strongest when the orientation of the stimulus in the RF matches the orientation of the surround (*39*). Tuned suppression is weakest for small image regions that differ in features from their surround that often belong to figures. It can therefore play a role in the segregation of small figures from the background. However, tuned suppression cannot explain FGM elicited by the phase-defined stimuli (*20*, *42*) described above (Fig. 2B) because the orientation of figure and ground is the same (Fig. 2B).

To further dissociate FGM from surround suppression, we created a stimulus in which the figure and background have different features that occupy the same surface area so that surround suppression is perfectly balanced. The figures were defined by convexity and closure, two of the Gestalt laws of perceptual organization (Fig. 7A). Human observers perceive the convex/enclosed regions as figures and the concave/nonenclosed regions as backgrounds (*43*–*45*), although perceptual organization can take hundreds of milliseconds for these more complex stimuli (*8*, *46*). We placed the stimuli so that the V1-RFs fell on either a convex/enclosed or a concave/nonenclosed region. The convex/enclosed regions elicited more V1 activity than the concave/nonenclosed regions, during a late phase of the response (given stimulus complexity, we chose an analysis window from 500 to 1000 ms, linear mixed-effects model, $P < 0.001$, 336 recording sites, 16 penetrations, eight mice; Fig. 7, B and C). Although the perceptual sensitivity of mice for convexity and closure remains to be determined, it is of interest that V1 activity is modulated by these Gestalt cues. The results, together, indicate that the enhanced activity elicited by figures goes beyond the release of surround suppression.

DISCUSSION

Our results provide new insights into how the interactions between V1 and HVAs shape perception. Upon appearance of a new image, the activity of V1 neurons exhibits a number of phases (*4*). During an early feedforward phase, activity is propagated from the retina via the thalamus to V1 and is transmitted onward to HVAs. V1 activity during this phase suffices for ceiling performance in the relatively simple contrast detection task, as if V1 acts as relay station that only needs to transmit information about the stimulus to HVAs (*47*). Inhibition of the entire visually driven V1 response reduced the accuracy of the mice. In accordance with a previous study (*48*), most animals were still able to perform the task above chance level, suggesting that information about the visual stimulus still reached the brain regions that are required for an appropriate licking response via alternative routes. Although we do not know what the mice perceived during V1 inhibition, comparable effects have been reported in humans with lesions in V1, who are able to correctly guess the presence, location, and shape of simple visual stimuli in the affected area of the visual field, despite denying any visual awareness of the stimuli (*49*). This phenomenon is known as blindsight and has been replicated in monkeys with a V1 lesion (*50*), where information from the lateral geniculate nucleus is directly propagated to HVAs, bypassing V1 (*51*).

If a visual stimulus requires figure-ground segregation, however, the peak response is followed by a phase in which the V1 neurons exhibit FGM; the representation of figural image elements is enhanced. In previous work, it remained unclear whether this late V1 activity phase is useful for perception, but our results demonstrate that optogenetic V1 silencing during this phase blocks figure-ground perception. This finding is in line with previous TMS studies in humans, although TMS produces much weaker perceptual effects (*52*). Among the studied figure-ground stimuli, we observed differences in the latency and strength of FGM, measured electrophysiologically (Fig. 2B), and in the V1 processing time required for figure-ground perception. FGM of orientation-defined figures occurred earlier and was stronger than FGM of phase-defined figures, and the mRT was also shorter for orientation-defined stimuli. Accordingly, V1 inhibition interfered with figure-ground perception at a later point in time for phase-defined than for orientation-defined figures.

The present results seem to differ from previous studies that examined the early and late phases of neuronal responses in the barrel

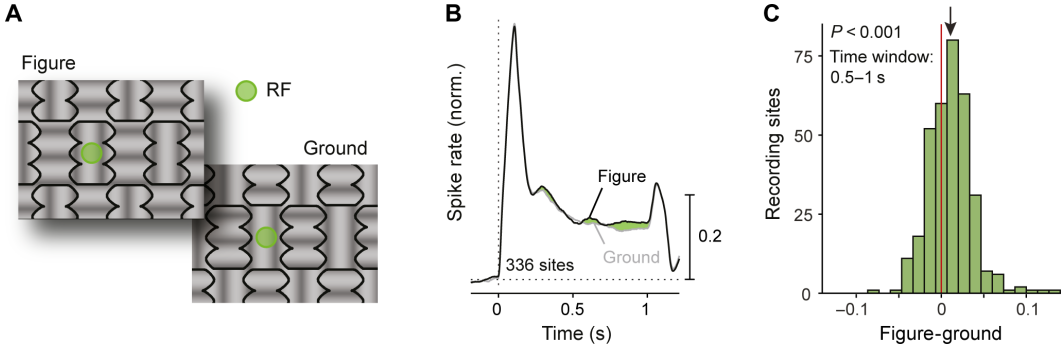

Fig. 7. The influence of convexity and closure on V1 activity. (**A**) Visual stimuli for which figure-ground perception depends on closure and convexity. Green circle, location of an example V1-RF in the putative figure and ground conditions. (**B**) Average activity across 336 recording sites in V1 (16 penetrations in eight mice) elicited by the convex/enclosed (black trace) or concave/nonenclosed image regions (gray trace). Green region, difference response. (**C**) Convex/enclosed regions elicit more activity than concave/nonenclosed regions (time window, 0.5 to 1 s, $P < 0.001$, linear mixed-effects model; see Materials and Methods).

cortex of mice engaged in tactile detection tasks (*53*, *54*). Suppression of late activity in the barrel cortex interfered with the detection of a tactile stimulus, whereas in the present study, late V1 activity was not required for the contrast detection task. We speculate that this difference between results may be related to the use of relatively weak tactile stimuli in these previous studies, which may have required amplification by recurrent interactions between lower and higher areas of the somatosensory cortex.

We identified a source for V1-FGM in HVAs that also exhibit FGM and project back to V1. Neurons in a number of HVAs exhibited FGM, which was strong in areas lateral from V1 (LM and AL) during passive viewing and weaker in anteromedial areas (RL, AM, and PM), although the wide-field experiments demonstrated that FGM increases in these areas when mice use the figure-ground stimuli to perform a task. The strength of FGM depends on the size of the RF relative to the figure size (*55*), but RF size differences between areas do not account for all interareal differences in FGM strength (fig. S5E). For example, the sizes of RFs in LM and RL are similar (*56*), but FGM is much stronger in LM, implying genuine differences in FGM amplitude across areas.

We also observed differences in the amplitude of FGM between the two-photon and wide-field imaging measurements, which presumably arise from different neuronal populations. In the two-photon experiments, we recorded activity from the cell bodies of L2/3 excitatory neurons, while wide-field calcium signals represent a mix of activity from laminar and cellular compartments with an emphasis on layer 1 (*30*). Understanding the functional organization of the mouse visual cortex is an active field of exploration (*26*, *57*, *58*), and the present results suggest that figure-ground segregation can help to further dissect the roles of higher cortical areas.

Optogenetic silencing of the HVAs strongly reduced FGM in V1, without a consistent effect on the early, visually driven response. This finding supports theories suggesting that FGM in V1 requires feedback from HVAs (*59*). The effect of silencing was pronounced for the phase- and texture-defined figure-ground stimuli and somewhat weaker if the figure was a grating with a different orientation. The weaker suppression is compatible with the view that horizontal interactions within V1 may also contribute to the perception of orientation-defined figures. Horizontal interactions within V1 are thought to be orientation selective so that suppression is stronger within image regions with a homogeneous orientation and weaker if a figure with one orientation is superimposed on a background with another orientation (*32*). If the figure is defined by a phase offset, the orientation selective suppression signal is absent, which may explain the larger dependence on corticocortical feedback. We note, however, that the viral construct was not taken up by all neurons in HVAs so that inhibition was presumably incomplete. Hence, we cannot exclude that the remaining FGM for orientation-defined figures reflects a feedback influence from the nonsilenced neurons in HVAs.

The main effect of silencing HVAs was a decrease in activity during the late phase of the V1-response, which was more pronounced for figures than the background, thereby decreasing FGM (fig. S6). These excitatory feedback effects imply that FGM is not caused by surround suppression and are consistent with predominantly excitatory (or disinhibitory) feedback effects on V1 activity (*60*, *61*). However, other studies in monkeys (*62*) and mice (*41*) suggested a predominantly suppressive influence of feedback connections. We identified a reason for this apparent discrepancy, because some of the previous studies on surround suppression focused on neurons with significant surround suppression (*41*, *63*), thereby emphasizing suppressive feedback influences, whereas the net feedback effect is excitatory when the entire population is taken into account (fig. S6, E and F). Other factors that may have contributed to these discrepancies include differences between the visual stimuli and the set of HVAs that were inhibited.

It has been suggested (*15*) that feedback connections enhance V1 activity at relevant locations by activating VIP interneurons, which, in turn, inhibit inhibitory SOM cells, thereby disinhibiting the cortical column. In accordance with this hypothesis, we observed that figures enhanced the activity of VIP neurons and suppressed the activity of SOM cells (Fig. 6A). Another previous study demonstrated that the activity of SOM-interneurons increases for larger stimuli and suggested that these cells contribute to surround suppression (*14*). Unlike this previous study, we here presented figure and ground stimuli with the same spatial extent. The stronger activity elicited by the background therefore implies that SOM cells do not integrate all image regions equally but that the increase in activity is most pronounced for stimuli with a homogeneous orientation. SOM cells were the only ones for which the ground elicited, on average, more activity than the figure, in accordance with the view that SOM cells have a role in orientation-tuned suppression. Furthermore, we found that inactivation of VIP neurons decreased V1 activity and diminished FGM, demonstrating that VIP-mediated disinhibition contributes to FGM. Our measurements of interneuron activity focused on the orientation defined figure-ground stimuli. We hypothesize that these interactions also generalize to other figure-ground display, but this hypothesis remains to be tested in future work.

PV-neurons were more active for figures than for the background, and this FGM signal was even stronger than that of pyramidal cells. Although the PV activity profile resembles that of VIP neurons, PV cells do not preferentially contact other inhibitory cell types to cause disinhibition. Instead, they suppress somatic regions of pyramidal neurons (*35*), making it unlikely that they are part of a disinhibitory circuit motif. PV cell activity was weak for the background, which, unlike for SOM cells, also rules out a specific role for PV-neurons in suppressing the pyramidal cell activity elicited by the background. Instead, PV-neurons have been shown to be involved in feedforward inhibition and controlling the gain of the cortical column (*64*, *65*), although they also receive direct input from feedback connections (*31*, *66*). The strong extra activity driven by figures is in accordance with the view that PV-neurons integrate the activity of nearby pyramidal neurons to control the gain of the cortical column. If pyramidal neurons become more active, the activity of nearby PV cells also increases, giving rise to a negative feedback loop (Fig. 6A). The high PV cell activity level elicited by figures is also in accordance with a study in the visual cortex in monkeys (*67*), which demonstrated that fast-spiking interneurons strongly increase their activity for behaviorally relevant stimuli.

The present results provide insight into how interactions between lower and higher areas of the visual cortex enable the coselection of image elements that belong to a single figure and their segregation from the background. These interactions are, in part, mediated by direct corticocortical connections, but there are also routes through subcortical structures [e.g., LP of the thalamus (*68*)], and future studies could compare the relative contributions of these different routes. As a result of these interactions, figural image elements

become labeled with enhanced neuronal activity across multiple areas of the visual cortex, in accordance with the view that such a labeling process determines the perceptual grouping of features of the same object (*5*, *69*, *70*). It seems likely that the findings generalize to more complex forms of perceptual organization. Imagine grasping an object that is surrounded by a number of other objects. The visual system guides our fingers to touch and grasp edges of the same object, a selection process (*71*) that has to rely on recurrent interactions between higher areas coding for the relevant object's shape and lower areas coding for its individual edges. Previous studies in monkeys demonstrated that the representation of a selected object's shape is enhanced in higher cortical areas (*72*) and that this also holds true for relevant edges in V1 (*69*). The present study illustrates how such selection process is coordinated across cortical areas. Future studies can now start to investigate the interactions between brain regions that are required for complex forms of perceptual organization and how they guide behavior (*5*).

MATERIALS AND METHODS

Experimental model and subject details

Thirty-five male and 14 female mice of 2 to 14 months age were used in this study. 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 J. Romaya at the LON (Laboratory of Neurobiology) at the Wellcome Department of Imaging Neuroscience) and linearized the luminance profile of the monitor/projector. Visual stimuli during passive electrophysiological experiments were projected onto a back-projection screen placed 15 cm from the mouse with a PLUS U2-X1130 DLP projector (mean luminance = 40.6 cd/m^2). The size of the projection was 76 cm by 56 cm, the field-of-view was 136° × 101.6°, the resolution was 1024 × 768 pixels, and the refresh rate was 60 Hz. For the behavioral electrophysiological experiments, the stimuli were presented on a 53.3 cm Liquid Crystal Display (LCD) screen with a resolution of 1280 × 720 pixels (Dell 059DJP) driven by a Windows computer at 60 Hz at a distance of 15 cm in front of the mouse. During the two-photon experiments, we presented visual stimuli to the left eye of the mice, using a 61 cm LCD monitor (Dell U2414H) with a resolution of 1920 × 1080 pixels and a refresh rate of 60 Hz, placed at an angle of 30° relative to the nose and a distance of 12 cm from the eye. For the optogenetic experiments, we used a 24-inch LCD monitor (1920 × 1200 pixels, Dell U2412M), placed 11 cm in front of the eyes and for wide-field experiments an LCD monitor (122 × 68 cm, Iiyama LE5564S-B1), at a distance of 14 cm. We applied a previously described (*73*) correction for the larger distance between the screen and the mouse at higher eccentricities. This method defines stimuli on a sphere and calculates the projection onto a flat surface. The orientation- and phase-defined figure-ground stimuli were composed of 100% contrast sinusoidal gratings with a spatial frequency of 0.075 cycles/deg and a mean luminance of 20 cd/m^2. The diameter of the figure was 35° (optogenetic- and wide-field imaging experiments), 40° (electrophysiology during passive viewing), or 50° (two-photon experiments). For the orientation-defined figures (Fig. 1, A and B), the grating orientation in the background was either horizontal or vertical, and the orientation of the figure was orthogonal. For the phase-defined figures (Fig. 1A), the phase of the figure grating was shifted by 180° relative to that of the background. For the contrast-defined stimuli, we presented the figure gratings on a gray background (20 cd/m^2). To test the generality of figure-ground perception, we also presented random textures in some of the experiments, again ensuring that the figure and background stimulus in the RF of the neurons was identical. The texture was made by filtering Gaussian distributed random noise patterns through an oriented filter. Four new random noise patterns were generated for each experimental session, and these were filtered with 0° and 90° filters, yielding eight oriented textures. The oriented filter (F) was made through summation of individual Gabor filters (G) as follows

$$F = \frac{\sum_{i=1}^{20} \frac{G_{f(i),\varphi}}{\mathrm{rms}\,(G_{f(i),\varphi})}}{20} \quad (1)$$

where G is Gabor filter of spatial frequency $f(i)$, where f was linearly spaced from 0.05 to 0.5 cycles/deg, orientation ϕ (0° or 90°), and SD *1/f*, RMS indicates a root mean square operation. After filtering, the texture was matched to the root mean square contrast of the sine-wave grating used in the other experiments using an iterative hard-clipping procedure with 10 iterations.

Surgical preparation

The mice were handled 5 to 10 min per day starting 1 week before surgery. Anesthesia was induced using 3 to 5% isoflurane in an induction box, and it was maintained using 1.2 to 2.5% isoflurane in an oxygen-enriched air (50% air and 50% O_2) mixture, and we subcutaneously injected 5 mg/kg meloxicam (0.5 mg/ml) as general analgesic. The mice were positioned in a stereotactic frame, and we monitored the depth of anesthesia by frequently checking paw reflexes and breathing rate. The temperature of the animal was monitored and kept between 36.5° and 37.5° with a heating pad coupled to a rectal thermometer. We covered the eyes with ointment to prevent dehydration. The area of incision was shaved and cleaned with betadine, and lidocaine spray was applied to the skin as a local analgesic.

Implantation of head fixation bar or ring

After induction of general anesthesia, we made an incision in the skin and cleaned the skull by removing any tissue and briefly applying H_2O_2. We then applied a dental primer (Keerhawe Optibond) to improve the bonding of cement to the skull. We fixed head bars or head rings in place using dental cement (Heraeus Charisma A1 or Vivadent Tetric Evoflow).

Virus injections

For the viral injections, we used a dental drill to make small craniotomies above the injection sites. We placed a pulled borosilicate capillary containing the virus vertically above the craniotomy touching the dura. Slowly the pipette was lowered to a depth of 600 µm from the brain surface, and we slowly injected the solution containing the virus at different depths using Nanoject III programmable injector (Drummond Scientific). After the injection, we left the pipette in place for at least 8 min before slowly retracting it to avoid efflux of virus.

Behavioral task

The mice were held on a reverse day-night cycle and a fluid restriction protocol with a minimal intake of 0.025 ml/g, while their health was carefully monitored. The animals were trained to indicate the side on which a figure appeared by licking the corresponding side of a custom-made double lick spout (Fig. 1B). We registered licks by measuring a change in capacitance with an Arduino and custom-written software. A trial started when the stimulus with a figure on the left or right appeared on the screen. The stimulus was displayed for 1.5 s. Because mice made early random licks, we disregarded licks from 0 to 200 ms (grace period). We prolonged the grace period to 500 ms for one mouse, because it helped correct a bias for preferentially licking one side (mouse M2 in table S1). The exact figure location varied slightly depending on the experiment, but the figure center was generally close to an azimuth of ±30° (left or right of the mouse) and an elevation of 15°.

Stimulus presentation was followed by a variable intertrial interval (ITI) of 6 to 8 s. Correct responses were rewarded with a drop of water or milk. If the animal made an error, a 5-s timeout was added to the ITI. We presented a background texture during the ITI and did not give reward if the mice licked so that they learned to ignore it. In some sessions, we included correction trials, which were repeats of the same trial after an error. We only included the noncorrection trials to compute accuracy, defined as hits/(hits + errors). During task performance in the wide-field imaging experiments, we used a motor which moved the lick spout close to the mouth of the mouse, 500 ms after the presentation of the stimulus, thereby enforcing a minimum viewing time before the mouse could respond. For the accuracies of individual mice, see table S1.

Electrophysiology

The laminar electrophysiological recordings during passive viewing (Fig. 2, A and B) were carried out in six Tg (Thy1-GCaMP6f) GP5.17Dkim mice (five males and one female) aged between 2 and 6 months. We implanted the mice with a head bar and one or two small screws over the cerebellum, which served as reference and ground and made a craniotomy centered on left V1 in a separate surgery preceding the first recording day. The mice were head-fixed and placed on a treadmill so that they could run or sit according to their preference. We tracked the treadmill movements using an Arduino and monitored pupil movements and size under infrared light with a zoom lens (M118-FM50, Tamron, Cologne, Germany) coupled to a camera (DALSA GENIE-HM640, Stemmer Imaging) and custom-written software. We inserted a linear-array recording electrode (A1x32-5 mm-25-177, NeuroNexus, 32-channel probe, 25-μm spacing) in V1 and lowered it to around 1 mm below the brain surface and adjusted the depth of the electrode with reference to the current source density profile as reported previously (*39*) to ensure coverage of all layers. We amplified the electrical signal from the electrodes and sampled it at 24.4 kHz using a Tucker Davis Technologies recording system. 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 multiunit spike rate. First, we measured the RF of the units recorded at each electrode with a sparse noise stimulus consisting of four white checks (8° by 8°, 40 cd/m^2) on a black background presented for 250 ms with a 250-ms ITI. The checks (>30 presentations per check) were positioned on a grid ranging from −64° to 16° horizontally and −22° to 66° vertically relative to the mouse's nose with negative values indicating the right hemifield. We corrected for flat screen distortion as was described above. We averaged the MUA response evoked by each check in a time window from 50 to 400 ms after check onset to obtain a map of visual responsiveness and fit a two-dimensional (2D)–Gaussian to estimate the width and center of the RF (Fig. 2A). The quality of the fit was assessed using r^2 and a bootstrapped variability index (BVI), which estimated the reliability of the RF center estimate. We resampled an equal number of trials as in the original dataset (with replacement) and regenerated the Gaussian fit. The BVI is defined as the ratio of the SD of the RF center position and the SD of the fitted Gaussian. To measure FGM, we centered a 40° diameter figure on the RFs. To create the background condition, we shifted the figure by 50° to 60° so that the RF fell on the background. As the RF position varies slightly across layers, we quantified the percentage of the RF area that fell within the figure boundary for each recording site. We calculated the overlap between the RF and the figure, such that at 100% overlap the RF was completely contained within the figure boundaries, and at 0% overlap, the RF was entirely outside the figure. We wished to exclude recording sites in which the RF contained a response to the boundary between the figure and the background. We empirically determined a cutoff value of overlap at which edge responses became apparent by examining the responses to 100% contrast gratings presented on a gray background. We binned the overlap values from all our recording sites into bins ranging from 0% overlap to 100% overlap. We observed a very sharp transition at 75% at which edge enhancement became apparent. We used a bin width of ±5% so we used 70% as a cutoff value. We only included recording sites if (i) the overlap between the RF and the figure was greater than 70% to exclude boundary-driven responses, (ii) the RF was reliable (r^2 of the Gaussian fit > 0.33, BVI < 1.5), (iii) if the signal-to-noise ratio of the visual response was greater than 1 [ratio of the activity between 0 and 100 ms after stimulus onset to the SD of baseline activity (−200 to 0 ms) across trials], and (iv) if the maximum response of the site was greater than 2 Hz. These criteria led to the inclusion of 198 recording sites for the electrophysiological data (Fig. 2B). The orientation-, phase-, or texture-defined figure-ground stimuli were presented in blocks of 32 stimuli. The position of the figure and the orientation of the underlying pattern/texture were pseudo-randomized within the block so that the RF stimulus was identical for the figure and ground conditions. To generate population responses, we subtracted the prestimulus activity (time window of −200 to 0 ms relative to stimulus onset) and normalized the activity at each recording site to the peak of the average, smoothed [Locally Weighted Scatterplot Smoothing (LOWESS) method, 39-ms window size] response across the figure and ground conditions (time window, 0 to 100 ms).

FGM was quantified with the d-prime, which is a measure for the reliability of the signal on individual trials (time window, 100 to 500 ms)

$$\text{d prime} = \frac{\mu_F - \mu_G}{\sqrt{\frac{1}{2}(\sigma_F^2 + \sigma_G^2)}} \quad (2)$$

where μ_F and μ_G are the means, and σ_F^2 and σ_G^2 are the variances of the figure and ground response across trials, respectively. To estimate whether the d-prime differed between cue types (i.e.,

orientation/phase/texture), we modeled the hierarchical correlation structure in the data with linear mixed-effects models to assess significance (*21*) using the fitlme.m function in MATLAB, because every electrode contained multiple contacts, and all contacts were tested for each of the cue types. Hence, there were two random effects: the recording contacts and the penetrations, which were included as random intercepts in the model. The two fixed effects were cue type (orientation-defined, phase-defined, and texture-defined figure) and the overall intercept (i.e., mean d-prime for the reference condition, which was one of the cue types). We obtained best model fits (as judged by a lower Akaike's Information Criteria) if we included a random slope term for cue type grouped by penetration. Differences between cue types were assessed by Bonferroni-corrected post hoc contrasts. To assess whether FGM was significant for each cue, we fit three separate linear mixed-effects models with a fixed intercept term and a random effect term for the penetration. As in previous studies, we estimated the latency of FGM by fitting a function (*12*) to the figure minus background response in a time window from 0 to 300 ms after stimulus onset. Briefly, the function is the sum of an exponentially modulated Gaussian and a cumulative Gaussian, capturing the Gaussian onset of neural modulation across trials/neurons and the dissipation of modulation over time. The latency (small arrows in Fig. 2B) was defined as the (arbitrary) point in time at which the fitted function reached 33% of its maximum value. Statistical comparisons of latency estimates were performed by bootstrapping. We selected a random set of recording sites (with replacement) 1000 times and recalculated the latency estimate for each resampled population.

Optogenetic silencing of V1 during task performance

For optogenetic silencing of V1 during behavior (Fig. 2, C to E), we injected five C57BL/6 mice (four males and one female) aged between 2 and 14 months with a cell-specific viral vector (AAV5-CaMKII-hGtACR2-eYFP) coding for the inhibitory GtACR2 opsin (*22*). The virus was cloned and produced in a custom preparation by Virovek Inc., Hayward, CA, USA based on the plasmid pFUGW-hGtACR2-EYFP, which was a gift from J. Spudich (Addgene plasmid #67877; http://n2t.net/addgene:67877 RRID: Addgene_67877). We injected a total of 150 nl per hemisphere (concentration of 3×10^{11} Genome copies (GC)/ml, three depths) in V1 at 2.7 mm lateral from the midline and 0.5 mm anterior of lambda. After 2 weeks, we attached a head fixation bar to the skull and applied a thin layer of adhesive to the bone that clarifies the skull, a method referred to as the "clear skull cap" technique (*28*). After the adhesive dried, we applied a thin layer of transparent dental cement (C&B Super-Bond) on top. We added a small rim of dental cement to the outer edges of the skull cap to prevent growth of the skin over the area of interest and applied a thin layer of transparent nail polish (Electron Microscopy Sciences) to reduce light glare.

Once the animals performed the task consistently with an accuracy larger than 65%, we introduced trials in which neural activity in the visual cortex was inhibited by activating the opsin. We presented figure-ground stimuli on 75% of all trials and applied optogenetic silencing in a random 25% of those trials. We presented contrast-defined stimuli in 25% of all trials, and because the accuracy was high and stable, we increased the fraction of trials with optogenetic intervention up to 50% (we used the same trial ratios in control experiments in which the laser was not directed at V1; fig. S3, D and E). We used a DPSS Laser (BL473T3-100FC, Shanghai Lasers & Optics Century Co.) emitting blue light (wavelength, 473 nm) as a light source and directed the light to the cortex with a split optical fiber [2 × 2 coupler, 50:50, 200 μm diameter, Ferrule Connector/Physical Connector (FC/PC) to 1.25-mm ceramic ferrule; Thorlabs, Newton]. The two fiber ends were directed at the centers of area V1 in the two hemispheres. Optogenetic stimulation lasted for 2 s with a constant power of 10 mW at each fiber end. The laser power at the skull cap was 5.6 mW/mm^2. The clear skull cap absorbs around 50% of the light (*28*) so that the effective laser power at the cortical surface was ~2.8 mW/mm^2, a light level that does not cause measurable heating (*74*). The onset of stimulation was shifted relative to the onset of the visual stimulus in steps of 16.7 ms according to the frame rate of the screen. The time between stimulus onset and laser onset was 17, 33, 50, 67, 83, 100, 150, or 200 ms. To remove spurious cues that might be caused by optogenetic stimulation, we placed a light shield around the head of the mouse to prevent light from the optogenetic stimulation reaching the eye and presented blue light flashes at random times by driving a blue LED placed below the mouse's head to flash every 0.5 to 1 s with an Arduino in all optogenetic experiments. We only included sessions in which the overall accuracy of the mice on laser off trials was above 70%. We used binomial tests (with Bonferroni correction) to examine whether performance was significantly above the chance level during trials with optogenetic manipulation and to test whether optogenetic inhibition impairs performance. To estimate the time at which the accuracy reached its half maximum, we fit a logistic function to the accuracy as a function of the laser onset latency using the Palamedes toolbox in MATLAB. We used bootstrapping (1000 times) by sampling trials with replacement to determine the 95% confidence interval of the latency, defined as the time when accuracy was halfway between the earliest and latest V1 silencing time point (i.e., the inflection point of the fitted function).

Wide-field imaging

In the wide-field imaging experiments (Fig. 3), we included eight males and one female Thy1-5.17 GCaMP6f mice (*27*) aged between 2 and 14 months. The skull of the animals was prepared with the clear skull cap technique (described above). We placed the mice under a wide-field fluorescence microscope (Axio Zoom.V16 Zeiss/Caenotec) to image a large part of the cortical surface. Images were captured at 20 Hz by a high-speed scientific Complementary metal–oxide–semiconductor (sCMOS) camera (pco.edge 5.5) and recorded using the Encephalos software package (Caenotec). We monitored the size and position of the right pupil (100-Hz sampling rate) and movements of the mouse with a piezo plate under the front paws of the mouse (100-Hz sampling rate) and removed trials with large body movements in a time window from the start of a trial until 350 ms after stimulus onset. We captured images of 1600 × 1600 pixels (~15 μm per pixel), down-sampled them to 400 × 400 pixels, and applied a Gaussian filter smoothing kernel of 5 × 5 pixels. We carried out a population RF mapping session [see below and (*75*)] to determine the borders of V1 and the HVAs, and we matched these areas to the Allen Brain common coordinate framework (fig. S5C). We computed the average $\Delta F/F$ (relative to the baseline fluorescence in a 300-ms window before stimulus onset) of all pixels within an area. In the active task, trials were removed if (i) the mouse had a performance below 65% in a window of 15 trials, or (ii) the absolute z-score of $\Delta F/F$ was larger than 3.5 (removal of trials with artifacts). We measured FGM with the d-prime in a time window of 150 to 300 ms after stimulus onset,

averaged across both hemispheres. We assessed significance with repeated measures ANOVAs, testing the effect of area (V1, LM, AL, RL, A, AM, PM, RSP, M1, and M2), figure-ground cue (orientation and phase defined), or response type (hit, error, and passive) on d-prime. We carried out the passive viewing experiments before the start of training in the figure-ground task ($N = 5$). See tables S1 and S2 for information about the mice in this experiment.

Two-photon imaging of excitatory neurons

We carried out experiments with five Thy1 mice aged 2 to 6 months old, including both sexes (see table S2) to image the activity of excitatory neurons (Fig. 4). The animals were anesthetized and implanted with a head ring as described above. After 2 weeks of recovery, we mapped the retinotopy [see below and (*75*) for the pRF mapping method] to determine the boundaries of visual areas. These mice were additionally injected with AAV1-CaMKII-GCaMP6f-WPRE-SV40 (Penn Vector Core, University of Pennsylvania, USA) in V1 (100 nl) and LM, AL, RL, AM, and PM (50 nl each) to enhance the GCaMP signal (concentration of 10^{12}GC/ml, injection speed of 20 nl/min, across three depths). To implant cranial windows, we made a circular craniotomy with a diameter of 4 to 5 mm centered at 0.5 mm anterior to lambda and 2.5 lateral from the midline. The craniotomy was closed with a double-layered glass coverslip, with the outer glass resting on the skull and fixed with dental cement (Vivadent Tetric Evoflow). After 2 weeks, we habituated the mice to head immobilization, while they could run on a running belt under the two-photon microscope (Neurolabware). We imaged through a 16× water immersion objective [Nikon, numerical aperture (NA) 0.80] at 1.7× zoom at a depth of 120 to 300 μm with a 15.7-Hz frame rate and a resolution of 512 x 764 pixels. We targeted V1 and other HVAs based on the retinotopic maps. A Ti-Sapphire femtosecond pulsed laser (MaiTai, Spectra Physics) was tuned to 920 nm for delivering excitation light. The power used varied between 20 and 50 mW depending on the depth of the imaging plane. First, we mapped the RF locations of the neurons within the field of view. We presented 12° × 12° white (38 cd/m^2) squares on a black background (0.05 cd/m^2) in an area ranging from −18° to 78° horizontally and −21° to 51° vertically relative to the mouse's nose. Each square was flashed twice for 166 ms with a blank interval of 166 ms, followed by delay of 500 ms. The square at each location was presented 20 times. We calculated RFs based on the response in a 500-ms window after stimulus onset. We fit a linear regression model to estimate the responses to the squares of the grid, 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 for every location to estimate the RF center and its full width at half-maximum response strength. We evaluated the quality of the fit using the r^2 value and the BVI (see above; r^2 of the Gaussian fit > 0.33, BVI < 1.5). We excluded cells with large RFs (SD of Gaussian fit > 20°) and which were not well-centered on the figure (Euclidean distance from RF center to figure center of >25°). The population of VIP cells had larger RFs, and so, we included any cell with an RF size of less than 40°. The figure-ground stimuli contained a 50° figure with one of six different orientations (maximum luminance of 38 cd/m^2 and minimum luminance of 0.05 cd/m^2). The figure was presented 120 times at two positions on the screen (20 repetitions per orientation), one centered on the RF location of a cluster of imaged cells, and the other at a distance of 55° from the RF center. The stimuli were presented in randomized order in blocks of 48 trials 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. We ensured that the stimulus in the RF was identical, on average, in the figure and background conditions, as described in the Introduction. We used CAIMAN (*76*, *77*) for preprocessing. We performed rigid motion correction for small shifts in the data due to 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 (*76*) 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. To generate population responses, we normalized the activity of each cell by subtracting baseline activity [time window, −325 to 0 ms (five frames) before stimulus onset] and dividing by the maximum of the average of the figure and ground conditions (time window, 0 to 2 s). For statistical analysis, we took the average baseline corrected $\Delta F/F$ values in a window from 0.3 to 1.5 s after stimulus onset for each cell in the figure and ground conditions. The distribution of these values was positively skewed, and we therefore took the natural log transform of the average $\Delta F/F$ values. Because the logarithm of negative values is undefined, we first removed cells with negative responses in either the figure or ground conditions. We removed outliers using an iterative multivariate outlier removal process. The Mahalanobis distance of each cell to the mean of the full distribution was calculated and z-scored. We removed cells for which the absolute z-score was greater than 3. If the maximum value of the preremoval z-score was greater than 6 (indicating the presence of an extreme outlier that may distort the calculation of the z-score), the process was repeated. This procedure removed less than 3% of cells. This outlier removal method was performed once for all the excitatory cell data from the different visual areas, after concatenating the data from each area. Similarly, the data from the different interneuron subclasses (described below) were concatenated together with the data from the excitatory cells in V1 before applying the outlier removal algorithm. The significance of the differences between figure and ground were first assessed using an omnibus linear mixed-effects model. Two models were analyzed, one for the data from the excitatory cells in different visual areas (i.e., the data shown in Fig. 4) and one for the data from different interneuron subclasses (including the excitatory cells in V1, i.e., the data in Fig. 6). The models contained two within-cell fixed effects terms {intercept [the mean log ($\Delta F/F$) in the background] and figure-ground condition} and one between-cell factor (either "visual area" or "cell class" depending on the model). The models also contained two random intercept terms for the cell identity and the imaging session in which the cell was recorded to account for any increased covariance between cells from the same imaging session. We also ran post hoc models for each visual area and cell class separately. These models contained two fixed effects (intercept and figure-ground condition) and two random intercept terms (cell identity and imaging session).

Electrophysiology with optogenetic inhibition of HVAs

The laminar electrophysiological recordings with optogenetic inhibition of HVAs (Fig. 5) were carried out in five male Tg (Thy1-GCaMP6f)GP5.17Dkim mice aged between 2 and 6 months. Surgical procedures were identical to those described above (under electrophysiology). We additionally applied a thin layer of

clarifying adhesive to the skull of the left hemisphere [clear skull cap (*28*)]. After 1 week of recovery, we used population RF mapping based on the GCaMPf expression (see pRF mapping below) to determine the borders between the visual areas. We targeted virus injections to HVAs based on these borders. We injected a total of 40 nl of AAV1-CaMKII-stGtACR2-FusionRed at each of the four locations, at a minimum distance of 500 μm from the V1 border (titer 1.5×10^{13}, across three depths). The construct encoding soma-targeted GtACR2 (*33*) (pAAV-CKIIa-stGtACR2-FusionRed) was a gift from O. Yizhar (Addgene viral prep #105669-AAV1; http://n2t.net/addgene:105669 RRID:Addgene_105669). We sealed the chamber with the biocompatible adhesive Kwik-Cast (World Precision Instruments). After 3 weeks, we made a craniotomy centered on the area of V1 with a population RF at 30° azimuth and 20° elevation. The mice were head-fixed on a treadmill, and we performed laminar recordings and RF mappings and placed figures as described above (under electrophysiology). We activated the laser on 50% of the trials, randomly interleaved with the trials without optogenetic inactivation. We used a DPSS Laser (BL473T3-100FC, Shanghai Lasers & Optics Century Co.) emitting blue light (wavelength, 473 nm) as light source and directed the light to lateral and medial HVAs through a split optical fiber (2 × 2 coupler, 50:50, 200-μm thickness, FC/PC to 1.25-mm ceramic ferrule; Thorlabs, Newton). We adjusted the power such that it equaled 5 mW at each fiber tip, resulting in an effective laser power of ~2.8 mW/mm^2 at the cortical surface, a light level that does not cause measurable heating (*74*). The laser was turned on 200 ms before onset of the visual stimulus and remained on for the entire duration of visual stimulation (500 ms) and was turned off 100 ms after stimulus offset. We normalized V1 activity to that in the figure and ground conditions without optogenetic intervention. We quantified FGM using d-prime as described above and removed extreme multivariate outliers (likely artifacts) by calculating the *z*-scored Mahalanobis distance of each recording site's d-prime$_{\text{laser off}}$ and d-prime$_{\text{laser on}}$ values from the full distribution. Recording sites with values of greater than 2.58 were removed (approximately 1% of recording sites). To estimate the significance of the laser-induced change in d-prime, we fit the data with a linear mixed-effects model containing two fixed effects (laser on/off and intercept, i.e., d-prime when the laser was off), two random intercept terms to account for covariability of data obtained from the same electrode contact site and penetration and one random-slope term to account for the fact that the effect of the laser varied across penetrations (fig. S6D). The same model was used to assess the laser-induced change in peak response (mean normalized activity in a window from 0 to 100 ms after stimulus onset). Laser-induced changes in figure and ground activity were assessed with a post hoc test including a random intercept term for electrode penetration.

We calculated a surround suppression index (SSix) for each unit (n) to estimate the strength of surround suppression in a subset of these experiments in which we also recorded responses to sine-wave gratings of 40° diameter centered on the RF (eight penetrations and 109 recording sites)

$$\text{SSix}(n) = \frac{\text{CENT}(n) - \text{SURR}(n)}{\text{SURR}(n)} \quad (3)$$

where CENT(n) is the average response of the unit to the 40° diameter grating (time window, 0 to 0.5 s) and SURR(n) is the average response of the unit to a full-screen grating (we used background of the phase-defined stimulus). Units that are surround suppressed will have weaker responses to the full-screen stimulus and hence a positive index.

We then examined the relation between this index and the effect of inhibiting the HVAs on activity in V1 in a separate subset of trials. The effect of HVA silencing on activity elicited by the background was quantified with a feedback silencing index (FBix). We used the background conditions from the orientation and phase-defined figures, to exclude figure enhancement influences, as follows

$$\text{FBix}(n) = \frac{\text{GND}_{\text{ON}}(n) - \text{GND}_{\text{OFF}}(n)}{\text{GND}_{\text{OFF}}(n)} \quad (4)$$

where GND_{ON} and GND_{OFF} are the average responses to the background of the orientation and phase-defined stimulus (time window, 0.1 to 0.5 s) when the HVAs were and were not inhibited, respectively. This index is positive if inhibiting the HVAs increases activity in V1. We assessed the significance of the relationship between SSix and FBix using a linear mixed effects model. We included two fixed effects (the intercept and slope of the relationship) and one random intercept term to account for covariability of data recorded from the same penetration.

Two-photon imaging of inhibitory neurons

We carried out experiments with four VIP-Cre, five SOM-Cre, and four PV-Cre animals between 2 and 6 months old including both sexes (Fig. 6, A to K, and table S2). The animals were surgically prepared as described above, and we injected a total of 200 to 300 nl of AAV9-CAG-flex-GCaMP6f-WPRE-SV40 (Penn Vector Core, University of Pennsylvania, USA) in right V1 (injection speed of 20 nl/min, across three depths). The craniotomy was sealed and the skin was sutured. Two weeks later, the animals underwent a second surgery to implant a head ring for immobilization and a cranial window to allow imaging of neuronal activity (see description of two-photon imaging method above). After 2 weeks, we habituated the mice to head immobilization while they could run on a running belt under the two-photon microscope. We mapped the RF of the neurons, presented figure-ground displays, and analyzed the data as described above.

Electrophysiology with optogenetic inhibition of VIP neurons

The laminar electrophysiological recordings with optogenetic inhibition of VIP neurons (Fig. 6, L to N) were carried out in three VIP-Cre mice (one male and two females) aged between 2 and 6 months. We targeted virus injections to left V1 (2.7 mm lateral from the midline, 0.5 mm anterior of lambda). We injected a total of 90 nl of AAV1-hSyn-SIO-stGtACR2-FusionRed (titer 1×10^{13} GC/ml, across three depths). The virus was a gift from O. Yizhar, Addgene viral prep #105677-AAV1. We allowed the virus to be expressed for 3 weeks before performing the electrophysiological experiments. To test the efficacy of the approach, we first inhibited VIP neurons in a condition with only the center grating in the RF and observed that silencing decreased visually evoked activity in V1 but that it did not influence spontaneous activity levels (fig. S7). In the main experiment with figure-ground stimuli, we activated the laser on 50% of the trials with a constant power of 2.83 mW/mm^2 directed at the V1 recording site through an optical fiber (200-μm thickness, FC/PC to 1.25-mm ceramic ferrule; Thorlabs, Newton). The laser was turned on 250 ms after onset of the visual stimulus and was turned off 100 ms after stimulus offset. We tested the

influence of optogenetic inhibition of VIP neurons on FGM in V1 using the same statistics as described in the section on the optogenetic inhibition of HVAs above.

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 post-fixated 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 (5× objective, Zeiss Plan-Apochromat, 0.16 NA) using custom-written Image-Pro Plus software and aligned the images to the Paxinos and Franklin adult mouse brain atlas (*78*). To determine the location of virus expression relative to the position of cortical visual areas, we imaged the intact ex vivo brains of M27-M31 with a wide-field microscope, averaging across 100 fluorescent images with an RFP filter using ThorCam (Thorlabs) software. To improve the signal-to-noise ratio, we normalized the contrast of individual images before averaging 100 images, corrected for unequal illumination by subtracting the blue channel from the red channel (FusionRed fluorescence), and smoothed the result with a 3 × 3 median filter. We used the population RF mapping data to determine the border of area V1 and used a bright-field image of the same field of view to visualize the blood vessel pattern and coregistered the loci of virus expression to the V1 border based on this blood vessel pattern (Fig. 5B).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/27/eabe1833/DC1

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

REFERENCES AND NOTES

1. R. Desimone, J. Duncan, Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
2. D. L. K. Yamins, J. J. DiCarlo, Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* **19**, 356–365 (2016).
3. K. Kar, J. Kubilius, K. Schmidt, E. B. Issa, J. J. DiCarlo, Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat. Neurosci.* **22**, 974–983 (2019).
4. 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).
5. P. R. Roelfsema, Cortical algorithms for perceptual grouping. *Annu. Rev. Neurosci.* **29**, 203–227 (2006).
6. V. A. Lamme, The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15**, 1605–1615 (1995).
7. J. Poort, M. W. Self, B. Van Vugt, H. Malkki, P. R. Roelfsema, Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex. *Cereb. Cortex* **26**, 3964–3976 (2016).
8. 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).
9. H. S. Scholte, J. Jolij, J. J. Fahrenfort, V. A. F. Lamme, Feedforward and recurrent processing in scene segmentation: Electroencephalography and functional magnetic resonance imaging. *J. Cogn. Neurosci.* **20**, 2097–2109 (2008).
10. 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).
11. P. C. Klink, B. Dagnino, M.-A. A. M.-A. Gariel-Mathis, P. R. Roelfsema, Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. *Neuron* **95**, 209–220.e3 (2017).
12. 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).
13. H. Supèr, H. Spekreijse, V. A. F. Lamme, Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat. Neurosci.* **4**, 304–310 (2001).
14. 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).
15. 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).
16. K. D. Harris, G. M. G. Shepherd, The neocortical circuit: Themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
17. 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).
18. M. M. Karnani, J. Jackson, I. Ayzenshtat, A. Hamzehei 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).
19. 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).
20. F. Li, W. Jiang, T.-Y. Wang, T. Xie, H. Yao, Phase-specific surround suppression in mouse primary visual cortex correlates with figure detection behavior based on phase discontinuity. *Neuroscience* **379**, 359–374 (2018).
21. H. Scheffe, A "mixed model" for the analysis of variance. *Ann. Math. Stat.* **27**, 23–36 (1956).
22. E. G. Govorunova, O. A. Sineshchekov, R. Janz, X. Liu, J. L. Spudich, Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science* **349**, 647–650 (2015).
23. A. Resulaj, S. Ruediger, S. R. Olsen, M. Scanziani, First spikes in visual cortex enable perceptual discrimination. *eLife* **7**, e34044 (2018).
24. H. Kirchner, S. J. Thorpe, Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vis. Res.* **46**, 1762–1776 (2006).
25. P. R. Roelfsema, F. P. de Lange, Early visual cortex as a multiscale cognitive blackboard. *Annu. Rev. Vis. Sci.* **2**, 131–151 (2016).
26. Q. Wang, E. Gao, A. Burkhalter, Gateways of ventral and dorsal streams in mouse visual cortex. *J. Neurosci.* **31**, 1905–1918 (2011).
27. 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).
28. Z. V. Guo, N. Li, D. Huber, E. Ophir, D. Gutnisky, J. T. Ting, G. Feng, K. Svoboda, Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**, 179–194 (2014).
29. G. N. Pho, M. J. Goard, J. Woodson, B. Crawford, M. Sur, Task-dependent representations of stimulus and choice in mouse parietal cortex. *Nat. Commun.* **9**, 2596 (2018).
30. W. E. Allen, I. V. Kauvar, M. Z. Chen, E. B. Richman, S. J. Yang, K. Chan, V. Gradinaru, B. E. Deverman, L. Luo, K. Deisseroth, Global representations of goal-directed behavior in distinct cell types of mouse neocortex. *Neuron* **94**, 891–907.e6 (2017).
31. W. Yang, Y. Carrasquillo, B. M. Hooks, J. M. Nerbonne, A. Burkhalter, Distinct balance of excitation and inhibition in an interareal feedforward and feedback circuit of mouse visual cortex. *J. Neurosci.* **33**, 17373–17384 (2013).
32. Z. Li, Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proc. Natl. Acad. Sci.* **96**, 10530–10535 (1999).
33. M. Mahn, L. Gibor, P. Patil, K. Cohen-Kashi Malina, S. Oring, Y. Printz, R. Levy, I. Lampl, O. Yizhar, High-efficiency optogenetic silencing with soma-targeted anion-conducting channelrhodopsins. *Nat. Commun.* **9**, 4125 (2018).
34. N. R. Wall, M. de la Parra, J. M. Sorokin, H. Taniguchi, Z. J. Huang, E. M. Callaway, Brain-wide maps of synaptic input to cortical interneurons. *J. Neurosci.* **36**, 4000–4009 (2016).
35. 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).
36. L. C. Garcia Del Molino, G. R. Yang, J. F. Mejias, X. J. Wang, Paradoxical response reversal of top- down modulation in cortical circuits with three interneuron types. *eLife* **6**, e29742 (2017).
37. C. M. Niell, M. P. Stryker, Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).
38. A. Vaiceliunaite, S. Erisken, F. Franzen, S. Katzner, L. Busse, Spatial integration in mouse primary visual cortex. *J. Neurophysiol.* **110**, 964–972 (2013).
39. 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).
40. L. Itti, C. Koch, Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 194–203 (2001).
41. 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).

42. D. S. Marcus, D. C. Van Essen, Scene segmentation and attention in primate cortical areas V1 and V2. *J. Neurophysiol.* **88**, 2648–2658 (2002).
43. M. Wertheimer, *A Brief History of Psychology* (Psychology Press, ed. 5, (2012); https://taylorfrancis.com/books/9781136620645.
44. M. Wertheimer, Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.* **4**, 301–350 (1923).
45. G. Kanizsa, W. Gerbino, in *Art and Artefacts* (NY: Springer, 1976), pp. 25–32.
46. D. Jeurissen, M. W. Self, P. R. Roelfsema, Serial grouping of 2D-image regions with object-based attention in humans. *eLife* **5**, e14320 (2016).
47. S. Thorpe, D. Fize, C. Marlot, Speed of processing in the human visual system. *Am J. Ophthalmol.* **122**, 608–609 (1996).
48. G. T. Prusky, R. M. Douglas, Characterization of mouse cortical spatial vision. *Vis. Res.* **44**, 3411–3418 (2004).
49. M. D. Sanders, E. K. Warrington, J. Marshall, L. Wieskrantz, "Blindsight": Vision in a field defect. *Lancet* **303**, 707–708 (1974).
50. A. Cowey, P. Stoerig, Blindsight in monkeys. *Nature* **373**, 247–249 (1995).
51. M. C. Schmid, S. W. Mrowka, J. Turchi, R. C. Saunders, M. Wilke, A. J. Peters, F. Q. Ye, D. A. Leopold, Blindsight depends on the lateral geniculate nucleus. *Nature* **466**, 373–377 (2010).
52. M. E. Wokke, I. G. Sligte, H. Steven Scholte, V. A. F. Lamme, Two critical periods in early visual cortex during figure-ground segregation. *Brain Behav.* **2**, 763–777 (2012).
53. S. Sachidhanandam, V. Sreenivasan, A. Kyriakatos, Y. Kremer, C. H. H. Petersen, Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat. Neurosci.* **16**, 1671–1677 (2013).
54. S. Manita, T. Suzuki, C. Homma, T. Matsumoto, M. Odagawa, K. Yamada, K. Ota, C. Matsubara, A. Inutsuka, M. Sato, M. Ohkura, A. Yamanaka, Y. Yanagawa, J. Nakai, Y. Hayashi, M. E. Larkum, M. Murayama, A top-down cortical circuit for accurate sensory perception. *Neuron* **86**, 1304–1316 (2015).
55. K. Zipser, V. A. F. Lamme, P. H. Schiller, Contextual modulation in primary visual cortex. *J. Neurosci.* **22**, 7376–7389 (1996).
56. J. H. Siegle, X. Jia, S. Durand, S. Gale, C. Bennett, N. Graddis, G. Heller, T. K. Ramirez, H. Choi, J. A. Luviano, P. A. Groblewski, R. Ahmed, A. Arkhipov, A. Bernard, Y. N. Billeh, D. Brown, M. A. Buice, N. Cain, S. Caldejon, L. Casal, A. Cho, M. Chvilicek, T. C. Cox, K. Dai, D. J. Denman, S. E. J. de Vries, R. Dietzman, L. Esposito, C. Farrell, D. Feng, J. Galbraith, M. Garrett, E. C. Gelfand, N. Hancock, J. A. Harris, R. Howard, B. Hu, R. Hytnen, R. Iyer, E. Jessett, K. Johnson, I. Kato, J. Kiggins, S. Lambert, J. Lecoq, P. Ledochowitsch, J. H. Lee, A. Leon, Y. Li, E. Liang, F. Long, K. Mace, J. Melchior, D. Millman, T. Mollenkopf, C. Nayan, L. Ng, K. Ngo, T. Nguyen, P. R. Nicovich, K. North, G. K. Ocker, D. Ollerenshaw, M. Oliver, M. Pachitariu, J. Perkins, M. Reding, D. Reid, M. Robertson, K. Ronellenfitch, S. Seid, C. Slaughterbeck, M. Stoecklin, D. Sullivan, B. Sutton, J. Swapp, C. Thompson, K. Turner, W. Wakeman, J. D. Whitesell, D. Williams, A. Williford, R. Young, H. Zeng, S. Naylor, J. W. Phillips, R. C. Reid, S. Mihalas, S. R. Olsen, C. Koch, A survey of spiking activity reveals a functional hierarchy of mouse corticothalamic visual areas. *bioRxiv* 10.1101/805010 (2019).
57. Q. Wang, O. Sporns, A. Burkhalter, Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. *J. Neurosci.* **32**, 4386–4399 (2012).
58. A. L. Juavinett, E. M. Callaway, Pattern and component motion responses in mouse visual cortical areas. *Curr. Biol.* **25**, 1759–1764 (2015).
59. P. R. Roelfsema, V. A. F. Lamme, H. Spekreijse, H. Bosch, Figure-ground segregation in a recurrent network architecture. *J. Cogn. Neurosci.* **14**, 525–537 (2002).
60. C. Y. L. Huh, J. P. Peach, C. Bennett, R. M. Vega, S. Hestrin, Feature-specific organization of feedback pathways in mouse visual cortex. *Curr. Biol.* **28**, 114–120.e5 (2018).
61. 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).
62. 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).
63. 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).
64. B. V. Atallah, W. Bruns, M. Carandini, M. Scanziani, Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73**, 159–170 (2012).
65. Y. Yoshimura, E. M. Callaway, Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nat. Neurosci.* **8**, 1552–1559 (2005).
66. R. D. D'Souza, A. M. Meier, P. Bista, Q. Wang, A. Burkhalter, Recruitment of inhibition and excitation across mouse visual cortex depends on the hierarchy of interconnecting areas. *eLife* **5**, e19332 (2016).
67. J. F. Mitchell, K. A. Sundberg, J. H. Reynolds, Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* **55**, 131–141 (2007).
68. S. M. Sherman, Thalamus plays a central role in ongoing cortical functioning. *Nat. Neurosci.* **19**, 533–541 (2016).
69. P. R. Roelfsema, R. Houtkamp, Incremental grouping of image elements in vision. *Atten. Percept. Psychophysiol.* **73**, 2542–2572 (2011).
70. K. M. O'Craven, P. E. Downing, N. Kanwisher, fMRI evidence for objects as the units of attentional selection. *Nature* **401**, 584–587 (1999).
71. D. Baldauf, H. Deubel, Attentional landscapes in reaching and grasping. *Vis. Res.* **50**, 999–1013 (2010).
72. L. Chelazzi, E. K. Miller, J. Duncan, R. Desimone, A neural basis for visual search in inferior temporal cortex. *Nature* **363**, 345–347 (1993).
73. J. H. Marshel, M. E. Garrett, I. Nauhaus, E. M. Callaway, Functional specialization of seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
74. S. F. Owen, M. H. Liu, A. C. Kreitzer, Thermal constraints on in vivo optogenetic manipulations. *Nat. Neurosci.* **22**, 1061–1065 (2019).
75. 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. *SSRN* (2020), doi:10.2139/ssrn.3676469.
76. 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).
77. 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).
78. G. Paxinos, K. B. J. Franklin, *Mouse Brain in Stereotaxic Coordinates* (ed. 2, 2004; doi: 10.1016/S0306-4530(03)00088-X).
79. 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).
80. S. O. Dumoulin, B. A. Wandell, Population receptive field estimates in human visual cortex. *NeuroImage* **39**, 647–660 (2008).
81. M. E. Garrett, I. Nauhaus, J. H. Marshel, E. M. Callaway, Topography and areal organization of mouse visual cortex. *J. Neurosci.* **34**, 12587–12600 (2014).
82. M. Sereno, A. Dale, J. Reppas, K. Kwong, J. Belliveau, T. Brady, B. Rosen, R. Tootell, Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* **268**, 889–893 (1995).

Acknowledgments: We thank the animal and mechatronics department at the Netherlands Institute for Neuroscience and I. Alonso, K. Koukoutselos, L. van Linge, T. van der Bijl, V. Bos, and A. Poerwoatmodjo for assistance in mouse training and electrophysiological recordings. We thank E. Ruimschotel for support with breeding and L. Cazemier for helpful comments. **Funding:** The work was supported by NWO (ALW grant 823-02-010) and the European Union's Horizon 2020 and FP7 Research and Innovation Program (grant agreements 720270 and 785907 "Human Brain Project SGA1 and SGA2", ERC grant agreement 339490 "Cortic_al_ gorithms," FLAG-ERA JTC grant ChampMouse and the Erasmus Mundus "NeuroTime" program) and the Friends Foundation of the Netherlands Institute for Neuroscience. **Author contributions:** U.H.S., E.H.v.B., and J.A.M.L. developed the task. L.K. and M.W.S. recorded and analyzed the electrophysiological data in which mice were passively viewing the stimuli. L.K. performed the V1 optogenetic silencing experiments during task performance. E.H.v.B. and A.B. performed the wide-field experiments, and E.H.v.B. and M.W.S. analyzed the results. S.M. carried out the two-photon experiments, and M.W.S., S.M., U.H.S., and C.v.d.T. analyzed the two-photon data. L.K. and M.W.S. recorded and analyzed the electrophysiological data in which we inhibited VIP neurons and the recordings in which we silenced HVAs. U.H.S. performed and analyzed the electrophysiological experiment during task performance. C.N.L. helped with the design of experiments involving transgenic mouse lines. P.R.R. and M.W.S. conceived of the experiments based on advice from J.A.H. and supervised the project. P.R.R., M.W.S., and L.K. wrote the paper. **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/KXJV-K22. Additional data related to this paper may be requested from the corresponding author.

Submitted 6 August 2020
Accepted 17 May 2021
Published 30 June 2021
10.1126/sciadv.abe1833

Citation: 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).

Science Advances

The essential role of recurrent processing for figure-ground perception in mice

Lisa Kirchberger, Sreedeep Mukherjee, Ulf H. Schnabel, Enny H. van Beest, Areg Barsegyan, Christiaan N. Levelt, J. Alexander Heimel, Jeannette A. M. Lorteije, Chris van der Togt, Matthew W. Self and Pieter R. Roelfsema

Sci Adv **7** (27), eabe1833.
DOI: 10.1126/sciadv.abe1833