



A central and unified role of corticocortical feedback in parsing visual scenes

Received: 24 September 2024

Accepted: 15 July 2025

Published online: 28 July 2025

 Check for updatesYe Xin^{1,2}, Yin Yan^{1,2,3,4}  & Wu Li^{1,2,4} 

Visual scene analysis in the brain involves diverse grouping and segmentation processes mediated by intra- and inter-areal interactions, but the role of feedback from higher to lower visual cortices has remained largely speculative, relying mostly on indirect evidence. In this study with behaving monkeys, we investigated the causal impact of V4-to-V1 feedback on different figure-ground perceptual tasks employing distinct contextual cues. We compared neuronal responses in the primary visual cortex (V1) and the behavioral performance before and after silencing a higher-order visual area (V4). Our results reveal that V4-to-V1 feedback comprises dissociable facilitatory and inhibitory components that differ in their spatial distribution, onset time, modulatory polarity, and magnitude. This feedback is essential for, and consistent across, distinct grouping and segmentation processes, operating independently of V1 neurons' selectivity for local features and primarily modifying the late phases of neuronal responses. These findings highlight fundamental differences between the mechanisms underlying global, feature-independent figure-ground organization and local, feature-dependent contextual analysis. Moreover, while the feedback modifies correlated variability in neural responses, this effect does not alter the neural population code. Our study underscores the central and unified role of corticocortical feedback in integrating and interpreting visual scenes, with implications for circuit-level mechanisms.

During visual scene processing, neurons in the primary visual cortex (V1) not only extract local visual features but also encode global contextual information, such as contour grouping based on the global continuity of line segments (Fig. 1a, top)¹ and surface segmentation driven by discontinuities in global features (Fig. 1a, bottom)^{2,3}. These grouping and segmentation processes have been linked to two complementary effects of contextual modulation observed in V1 of monkeys: facilitation of neurons with receptive fields (RFs) falling on the foreground (contour or surface) followed by inhibition of neurons with RFs within the background areas adjacent to foreground elements^{4–7}. These two effects, both of which are significantly delayed relative to stimulus-evoked V1 responses but earlier than top-down attentional modulation of the figure-ground signals^{3,7}, are attributed to feedback influences from

higher visual cortex. Previous studies in monkeys^{6,8–14} and rodents^{15–19} have provided causal evidence for corticocortical feedback influences on V1 neuronal responses. However, it remains elusive whether these documented feedback modulations directly underlie the delayed foreground facilitation and the further delayed background inhibition.

On the other hand, long-range horizontal connections within V1^{20–22} are thought to be involved in feature-dependent contextual effects and center-surround modulations of neuronal responses (for a review see ref. 23). For instance, the response of an orientation-selective neuron to a bar centered in its RF is influenced not only by the bar's orientation but also by the arrangement of surrounding bars^{24–26}. In particular, V1 neurons can encode the orientation contrast between a bar and surrounding iso-oriented bars (orientation singleton; Fig. 1a,

¹State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China. ²IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. ³Chinese Institute for Brain Research, Beijing, China. ⁴College of Life Sciences, Beijing Normal University, Beijing, China.

✉ e-mail: yanyin@bnu.edu.cn; liwu@bnu.edu.cn

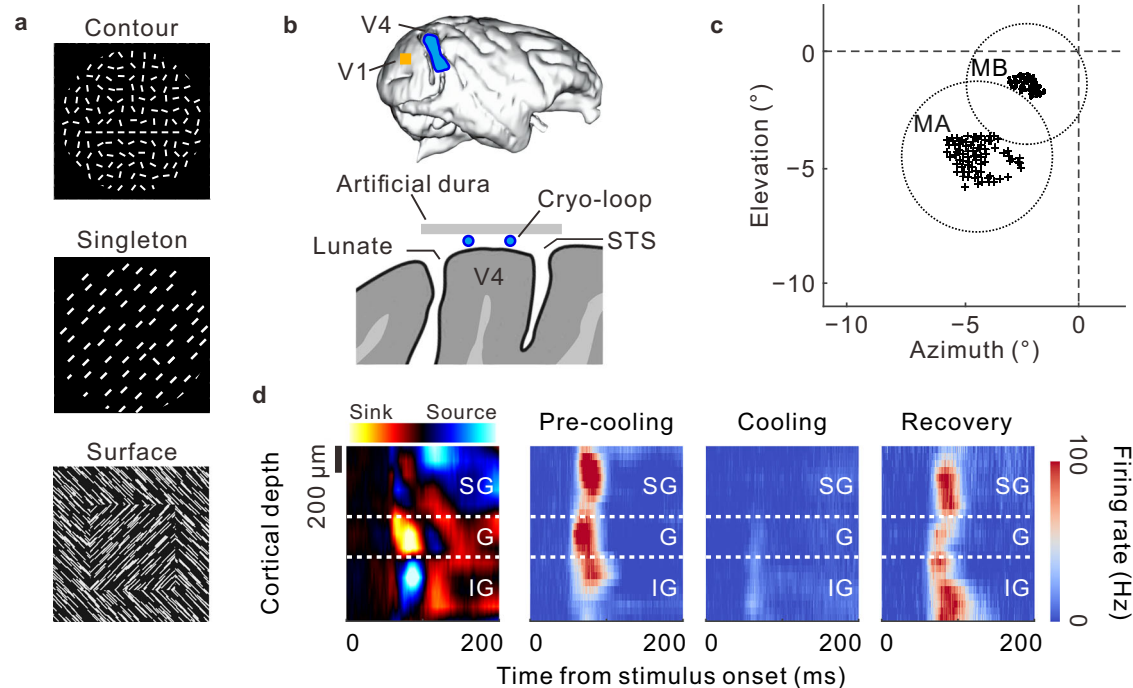


Fig. 1 | V1 recording and V4 inactivation. **a** Contour, orientation singleton, and surface stimuli commonly used for studying contextual influences. **b** Diagram of V1 recording and V4 inactivation. A cryo-loop was used to maintain $2.2^{\circ}\text{C} \pm 0.3^{\circ}\text{C}$ surface temperature of the V4 region (blue) between the lunate and superior temporal sulcus (STS; see also Supplementary Fig. S1a, c). A microelectrode array was implanted in the retinotopically matched V1 region (yellow). **c** Distribution of V1 receptive fields (RFs) recorded from the two monkeys (MA and MB; see also Supplementary Fig. S1b, d). Each data point represents the RF center of a V1 site recorded by an electrode. Circle illustrates the visual-field area covered by the stimuli, 6.5° in diameter for MA centered

at $(-4.4^{\circ}, -4.5^{\circ})$; 5.2° in diameter for MB centered at $(-2.2^{\circ}, -1.4^{\circ})$. **d** Validation of the effectiveness and reversibility of V4 cooling. A linear multi-contact probe (24 recording sites spaced $100\text{ }\mu\text{m}$ apart) was inserted into V4 perpendicular to the cortical surface. Horizontal dashed lines delimit the supra-granular (SG), granular (G), and infra-granular (IG) layers, as estimated by the current source density analysis (left panel; Methods). Compared to the pre-cooling and recovery stages, cooling effectively silenced V4 neuronal spiking activity evoked by gratings. Data from 4 penetrations in MB were aligned, averaged, and smoothed via interpolation across cortical depth. Source data are provided as a Source Data file.

middle) even in anaesthetized animals²⁷. Unlike the delayed signals for global contours and surfaces, local orientation contrast signals appear early in V1 responses, suggesting bottom-up feedforward processing²⁸. This is in line with behavioral observations that stimulus-driven saliency effects caused by orientation singletons are short-lived, emerging immediately and briefly after stimulus onset²⁹. Despite their early emergence, these orientation-dependent contextual signals are relatively weak but are significantly amplified in later V1 responses when monkeys report the presence of the singleton²⁶. Although various figure-ground cues (global contours, local singletons, and surfaces; Fig. 1a) are linked to similar late V1 responses, it is not yet clear whether these different grouping and segmentation processes share a common feedback modulatory mechanism. Additionally, the relationship between the mechanisms underlying contextual modulations in figure-ground organization and those involved in feature-dependent center-surround interactions remain to be determined.

In this study, we aim to address the following questions: (1) the necessity of V4 feedback for both foreground facilitation and background inhibition, (2) the separability of feedback components mediating these effects, (3) the generality of feedback modulation across different grouping and segmentation processes, and (4) the relationship between contextual modulations in figure-ground organization and center-surround interactions. To explore these questions, we compared V1 neuronal responses in monkeys before and after inactivating V4, a mid-level visual area involved in processing global forms³⁰ and providing direct and indirect feedback connections to V1^{31–33}.

Results

Responses of superficial layer V1 neurons were recorded in two macaque monkeys (named MA and MB) using a 10×10 microelectrode

array implanted in the right hemisphere. Reversible inactivation of V4 was achieved with a customized cooling setup³⁴ (Fig. 1b, Methods). The inactivated V4 region and the recorded V1 region were retinotopically matched to the visual-field area where stimuli were presented (lower left visual field, Fig. 1c; see also Supplementary Fig. S1a–d). The effectiveness and reversibility of the cooling were validated using a linear probe, allowing simultaneous recordings of spiking activities across V4 layers (Fig. 1d).

Impact of V4 inactivation on behavior

We trained the monkeys using a Go/No-go procedure under time pressure to detect whether a visual contour was present against a background of randomly oriented bars (Fig. 2a). In 50% of the trials (the Go trials, Fig. 2a, upper row), a contour pattern was displayed. The contour was formed by collinearly aligning 3, 5, 7, or 9 adjacent bars in a row, offset by one row from the pattern center. In the remaining 50% of the trials (the No-go trials, Fig. 2a, bottom row), a noise pattern with no contour was shown. After stimulus onset, the monkeys were required to make a saccade to the contour pattern within 500 ms or to maintain fixation for 500 ms for the noise pattern.

In the cooling experiment conducted on a given day, the monkeys completed the contour detection task in two successive sessions: a pre-cooling session with intact V4 followed by a cooling session with V4 inactivated. Each session included 200 Go trials randomly interleaved with 200 No-go trials. To control for potential confounding effects of adaptation and fatigue, we conducted a control experiment on a separate day. This control experiment was identical to the cooling experiment, except that V4 remained intact during both sessions (referred to as the pre-sham and sham-cooling sessions, respectively). To test a range of contour orientations and positions relative to the

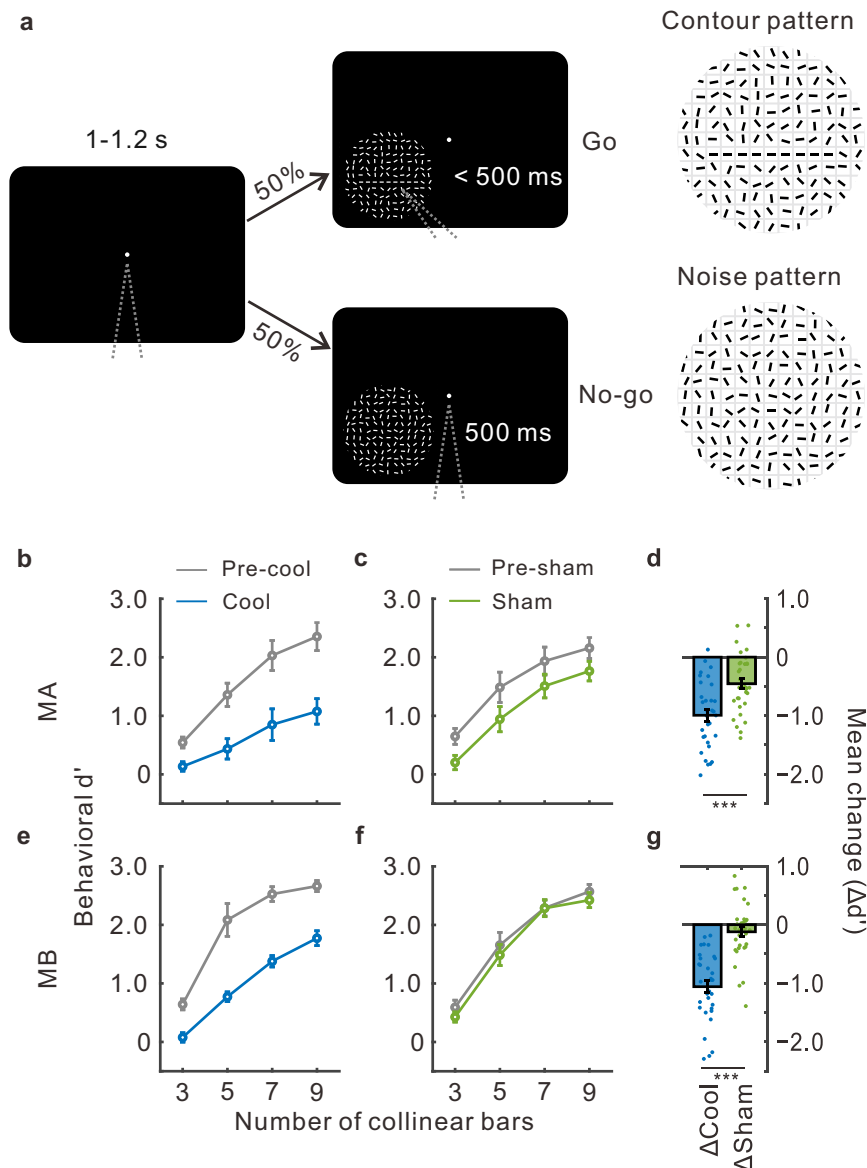


Fig. 2 | Impaired contour detection performance after V4 inactivation.

a Contour detection task. The monkey fixated on a small spot (left) until the onset of the stimulus pattern (middle and right), which consisted of small bars ($0.25^\circ \times 0.05^\circ$ for MA, $0.2^\circ \times 0.04^\circ$ for MB) distributed in a grid of 13×13 invisible square compartments (0.5° for MA, 0.4° for MB). The stimulus included randomly oriented background bars, with a contour formed by collinear bars either present (upper row) or absent (lower row). The monkey was required to make a saccade to the contour within 500 ms (Go trials) or maintain fixation (No-go trials). The contour was offset by one row in the invisible grid (upper right), so rotating the circular pattern around its center varied both the position and orientation of the contour. The contour position was fixed for each day, and a total of 8 contour positions were tested on separate days by rotating the stimulus in multiples of 45° . For each

position, the contour in the Go trials could be one of 4 lengths (3, 5, 7 or 9 collinear bars) with equal probability. **b–d** Behavioral results from monkey MA. Contour detection performance (behavioral d') is plotted as a function of contour length in the cooling experiment (**b**) and the sham-cooling control experiment (**c**). Each data point on the curve represents the average across the 8 contour positions (i.e., $n = 8$; see **a** legend). The overall changes in d' after V4 cooling ($\Delta Cool$) and sham-cooling ($\Delta Sham$) were calculated, respectively, relative to the pre-cooling and pre-sham sessions by averaging across the 4 contour lengths and 8 contour positions (**d**, $n = 32$, $p = 3.0 \times 10^{-5}$, two-tailed Wilcoxon signed-rank test). **e–g** Similar results from monkey MB (**e**, $n = 8$; **f**, $n = 8$; **g**, $n = 32$, $p = 1.4 \times 10^{-6}$, two-tailed Wilcoxon signed-rank test). Error bars represent SEM; ***, $p < 0.001$. Source data are provided as a Source Data file.

cluster of recorded V1 RFs, we rotated the entire stimulus pattern around its center in multiples of 45° , generating a total of 8 possible contour positions along a circular path. These different contour positions were tested on different days.

In the pre-cooling session (of the cooling experiment) or the pre-sham session (of the control experiment), the animals' behavioral sensitivity to perceiving the contour (measured as behavioral d' ; Methods) increased with contour length (Fig. 2b, c for monkey MA, Fig. 2e, f for MB, gray curves). In the cooling session following V4 inactivation, behavioral performance was dramatically impaired

(Fig. 2b, e, blue vs. gray). Extended experimental sessions per se also caused a general decline in performance, likely due to fatigue or adaptation, as observed in the control experiment (Fig. 2c, f, green vs. gray; more evident in Fig. 2c for MA). Nonetheless, the adaptation or sham-cooling effect averaged across the tested stimulus conditions (4 contour lengths and 8 contour positions) was much smaller than the effect of V4 cooling (Fig. 2d for MA, green vs. blue, $p = 3.0 \times 10^{-5}$; Fig. 2g for MB, $p = 1.4 \times 10^{-6}$; two-tailed Wilcoxon signed-rank test, $n = 32$). Based on the cooling and sham-cooling effects, we computed the net percent decrease in behavioral d' solely due to V4 inactivation

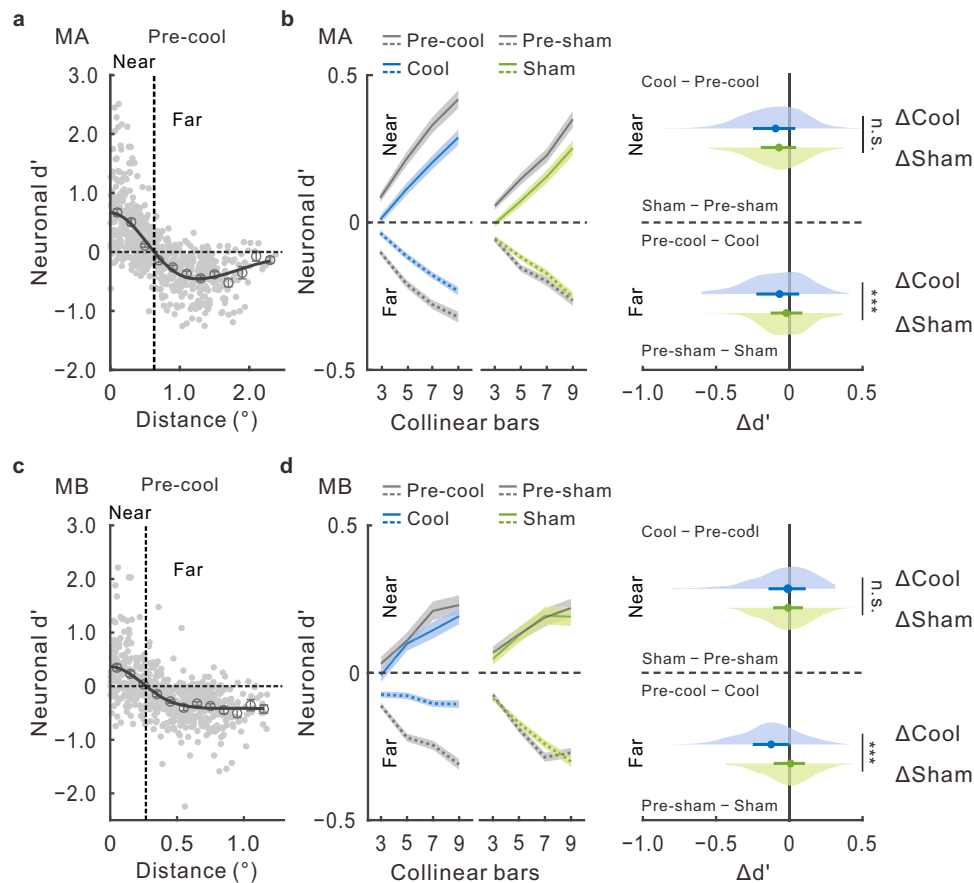


Fig. 3 | Impact of V4 inactivation on contour-induced modulations in V1.

a Spatial profile of contour-induced modulations in V1 of monkey MA. Gray dots represent individual V1 sites, with pre-cooling d' values (9-bar contour vs. noise pattern) plotted against distance from the RF center to the contour path ($n = 681$ sites pooled across 8 contour positions). Binned averages (open circles) were fitted with a difference-of-Gaussians (DoG) curve. Its intersection with the horizontal dashed line ($d' = 0$) defined the boundary between near and far site groups (vertical dashed line). **b** Comparison of V4-cooling and sham-cooling effects on V1 near and far sites in MA. Mean neuronal d' is shown as a function of contour length for cooling (left column: near sites, $n = 340$; far sites, $n = 341$) and sham-cooling (middle column: near sites, $n = 328$; far sites, $n = 335$; light shading: SEM). For each site group, the reduction in contour modulation strength ($\Delta d'$) was computed by pooling data across contour

lengths. Half-violin plots (right column) show distributions of $\Delta d'$ for cooling (ΔCool) and sham-cooling (ΔSham), with medians (dots) and interquartile ranges (horizontal colored lines). Median ΔCool and ΔSham values were compared for near sites (upper panel; $p = 0.054$) and far sites (lower panel; $p = 4.5 \times 10^{-5}$, two-tailed Wilcoxon rank-sum test). **c** As in **a**, but for monkey MB ($n = 619$). Compared with MA, the spatial scale of contour-induced modulations differed substantially (note x-axis range), likely due to differences in RF eccentricities between animals (see Fig. 1c). **d** Similar to **b**, but for MB. Left column: cooling experiment (near sites, $n = 211$; far sites, $n = 408$). Middle column: sham-cooling (near sites, $n = 248$; far sites, $n = 372$). Right column: $\Delta d'$ distributions (upper panel, $p = 0.053$; lower panel, $p < 10^{-6}$). Significance: n.s., $p \geq 0.05$; ***, $p < 0.001$. Source data are provided as a Source Data file.

(Methods, Data analyses). The net cooling effect was comparable between the two animals (42.1% for MA and 46.7% for MB). The impact of V4 cooling on contour detection is consistent with previous V4 lesion studies, which have reported severe impairments in form perception^{35–37}.

During the contour detection task in the cooling and control experiments, we simultaneously recorded V1 responses.

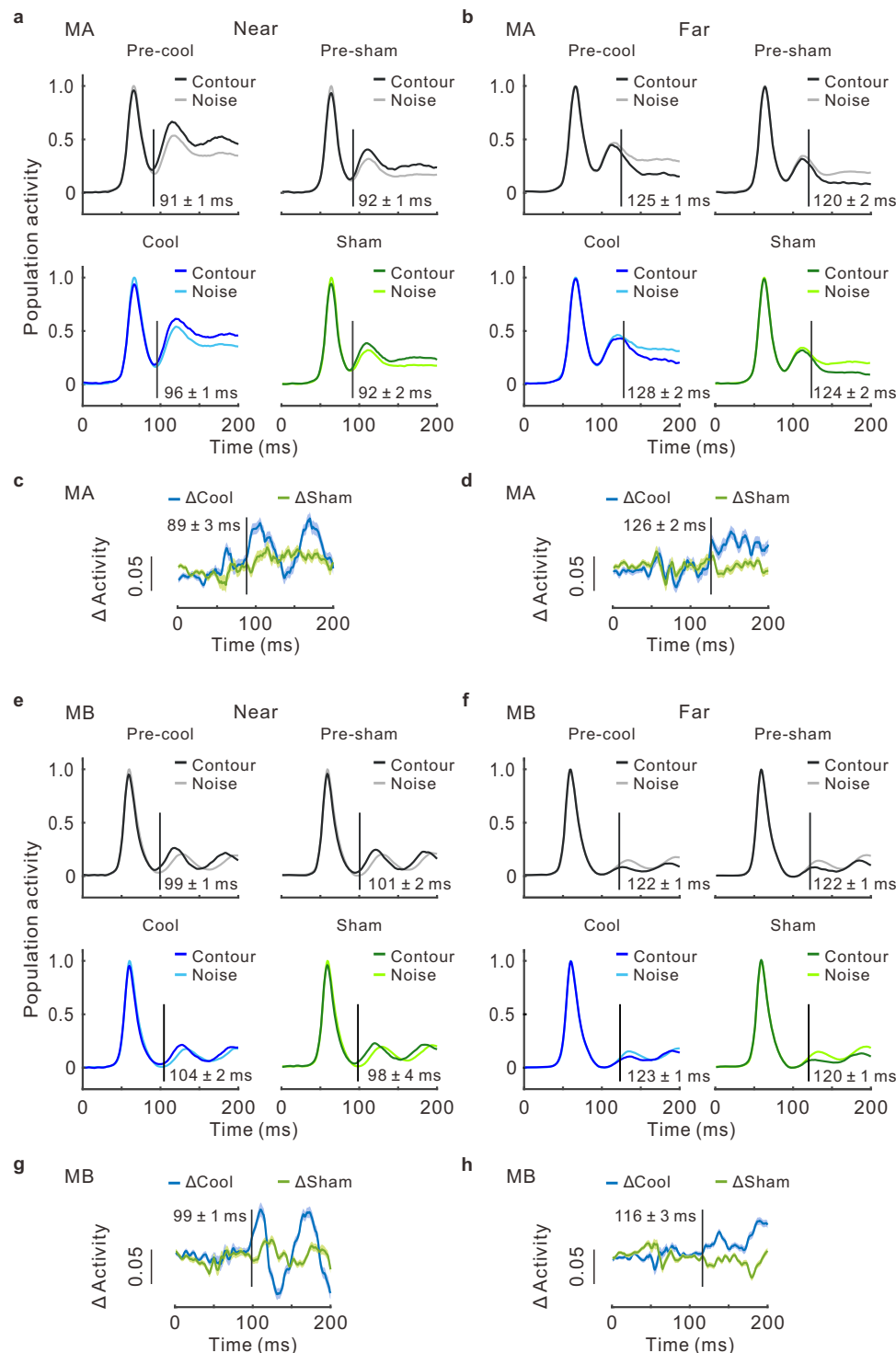
Dissociable modulatory effects of V4-to-V1 feedback

Previous studies have demonstrated two complementary modulatory effects in V1 induced by global contours: contour facilitation for neurons with RFs located on the contour and background inhibition for neurons with RFs located in the background, creating a center-surround profile across V1 neurons^{4,7,38}. These contour-induced modulatory effects were replicated during the pre-cooling and pre-sham sessions with intact V4.

The sensitivity of a V1 site to differentiate between contour and noise patterns was quantified as the neuronal d' (Methods), based on trial-by-trial spike counts recorded 90–200 ms after stimulus onset. To map the overall spatial profile of contour-induced modulations across

V1 sites, the d' values (9-bar contour pattern vs. noise pattern) from all sites were plotted against the distances between the RF centers and the collinear contour (Fig. 3a, c for the two animals during pre-cooling sessions; similar results were obtained from pre-sham sessions). Since contour-induced modulations depended on the RF-to-contour distances, data collected from the same electrode for different contour positions were taken as different samples. The scatter plot was fitted with a difference of Gaussians (DoG) curve, separating V1 sites into two groups: near sites and far sites, representing RFs located on the contour and in the background, respectively. On average, these two groups showed opposite effects: contour facilitation for near sites and background inhibition for far sites, indicated by positive and negative d' values. Both effects increased with contour length (Fig. 3b, d, solid and dotted gray curves; left column, pre-cooling session; middle, pre-sham session).

To isolate the true impact of V4 inactivation on V1 neuronal d' while controlling for potential adaptation effects (see Fig. 2c, f), we compared data from V4 cooling and sham-cooling sessions with their respective pre-cooling and pre-sham sessions (Fig. 3b, d, left and middle columns, paired color and gray curves). A qualitative inspection revealed that adaptation in V1 was more pronounced in MA than in



MB across contour lengths (Fig. 3b vs. 3d, middle column, green vs. gray), consistent with behavioral observations (Fig. 2c vs. 2f). After pooling different contour lengths and quantitatively comparing the overall effects of cooling and sham cooling, we observed consistent results for both MA and MB (Fig. 3b, d, right column, half-violin plots). V4 cooling caused a significantly greater reduction in background inhibition compared to sham cooling (lower panels, Δ Cool vs. Δ Sham, $p = 4.5 \times 10^{-5}$ for MA, $p < 10^{-6}$ for MB, two-tailed Wilcoxon rank-sum test) but did not significantly affect contour facilitation (upper panels, Δ Cool vs. Δ Sham, $p = 0.054$ for MA, $p = 0.053$ for MB). This differential impact became more apparent when calculating the net cooling effect, with the sham-cooling effect deducted (Methods, Data analyses). The

net percent decrease in background inhibition was 26.4% for MA and 56.2% for MB, while the net decrease in contour facilitation was only 2.4% and 18.2%, respectively. These analyses suggest that the effects of V4 inactivation—and thus the influence of V4-to-V1 feedback—were asymmetric, with a more pronounced impact on background inhibition than on contour facilitation. This finding was further supported by a linear mixed-effects model analysis accounting for potential repeated sampling from individual electrodes (Supplementary Table S1).

A comparison between the population-averaged post-stimulus time histograms (PSTHs; Methods) in response to the longest 9-bar contour pattern and the noise pattern revealed that both the contour facilitation of near sites (Fig. 4a, e for the two monkeys) and the

Fig. 4 | Time course of contour-induced modulations in V1 before and during V4 cooling. a–d Data from monkey MA. **a** Contour facilitation. Normalized population post-stimulus time histograms (PSTHs; Methods) for V1 near sites, showing responses to the longest, 9-bar contour pattern (darker curves) and the noise pattern (lighter curves) during the cooling experiment (left column) and the sham-cooling experiment (right column). The vertical line marks the contour facilitation latency, indicating where the paired curves start to diverge in the late neural responses. Latency was estimated using a curve-fitting approach (Methods) applied to the differential PSTH (contour – noise) of each pair of PSTHs, and the mean \pm SEM was calculated by resampling the near sites ($n = 340$ for the cooling experiment; $n = 328$ for the sham-cooling experiment). **b** Background inhibition. Similar to **a**, but with differential PSTHs derived from (noise – contour) for estimating the latency of background inhibition ($n = 341$ for the cooling experiment; $n = 335$ for the sham-cooling experiment). **c** Feedback influence on contour

background inhibition of far sites (Fig. 4b, f) were confined to the late components of V1 responses. To infer the timing of feedback modulation, we quantified the time course of contour-induced modulations during the cooling and sham-cooling experiments using a curve-fitting approach³⁹ (Methods). Contour facilitation consistently preceded background inhibition by approximately 20 ms (Fig. 4a vs. 4b for MA; 4e vs. 4f for MB). These latencies were minimally affected by V4 cooling or sham cooling (compare top and bottom panels in Fig. 4a, b, e, f). Notably, the effects of V4 cooling and sham cooling began to diverge for V1 near and far sites with latencies similar to those of contour facilitation and background inhibition (Fig. 4c, d for MA; Fig. 4g, h for MB), indicating a tight coupling between feedback influences from V4 and contour-induced modulations in V1.

Similar to the 9-bar contour, comparable latency values were obtained for a shorter 5-bar contour (Supplementary Fig. S2). These results are consistent with a previous observation in monkeys with intact V4: the onset time of contour-induced modulation in V1 remains unchanged with contour length⁴. As these modulatory effects occurred no earlier than 90 ms after stimulus onset, we chose this time point to separate V1 responses into early (0–90 ms) and late (90–200 ms) epochs for all relevant analyses. The 200 ms cut-off was determined based on behavioral reaction times (Supplementary Fig. S3).

The results from V4 inactivation suggest that the two complementary modulatory effects observed in V1 during contour grouping are closely tied to V4-to-V1 feedback influences, which are dissociable in terms of space, time, polarity, and magnitude.

Feedback influences on neural population code

The cooling experiment suggests that V4-to-V1 feedback enhances contour grouping and background segregation by modulating V1 neuronal firing rates. However, the amount of information conveyed by the neural population may also be affected by trial-by-trial fluctuations in individual neuronal activity (i.e., response variance) and by the structure of correlated variability among neurons (i.e., noise correlation, r_{sc})⁴⁰. Focusing on the late response epoch (90–200 ms), we first examined the effects of V4 cooling on pairwise r_{sc} by combining V1 sites into near-near, far-far, and near-far groups. This analysis showed differential impacts of V4 cooling on the r_{sc} across the three groups of paired V1 sites (see Supplementary Fig. S4 and Table S2). By pooling both near and far sites, we next examined whether feedback modulation could enhance the population code for differentiating between contour and noise patterns by altering variables other than firing rate.

Neural population coding was analyzed using Fisher's Linear Discriminant (FLD), which accounted for the correlated variability among all V1 sites and estimated the contour signal strength carried by the entire neural population. In this analysis (Methods), trial-by-trial population response vectors of all recorded V1 sites were projected onto the FLD axis. This axis optimally separated the projected population responses to a given contour pattern from those to the noise pattern, allowing us to compute a population d' as a measure of the contour signal strength

conveyed by population activity. Consistent with the analysis of individual V1 sites (see Fig. 3b, d), V4 cooling resulted in a significantly greater reduction in V1 population d' compared to sham cooling (Fig. 5, 1st column, monkey MA; Fig. 5b, 1st column, MB). Neural population activity can be influenced by three variables: mean firing rates, response variances, and the r_{sc} structure. To isolate the components of feedback-induced changes in neural population code, we constructed surrogate V1 datasets with selectively manipulated variables. The validity of this approach was confirmed as follows: For each stimulus condition in each experimental session, we used trial-averaged firing rates and the covariance matrix of individual V1 sites to construct a multivariate normal distribution. This multivariate Gaussian model, uniquely defined by its means, variances, and covariances, provided a comprehensive representation of the data. This distribution was used to generate simulated trial-by-trial spike counts, preserving the key metrics of the original dataset and replicating the FLD analysis results (Fig. 5a, b, 2nd column, nearly identical to the 1st column). After validating the method, we conducted the following three simulations to isolate the effects of different variables.

The first simulation examined the effect of r_{sc} structure change. We synthesized a covariance matrix incorporating pre-cooling response variances and post-cooling Pearson correlation coefficients. With this covariance matrix and pre-cooling mean firing rates, we created a surrogate cooling dataset reflecting the post-cooling r_{sc} structure, while preserving pre-cooling means and variances. A similar surrogate sham-cooling dataset was created for comparison. FLD analysis showed that changes in the r_{sc} structure alone due to V4 inactivation did not decrease population d' ; in fact, a slight increase was found (Fig. 5a, b, 3rd column; blue vs. green bar, $p = 6.3 \times 10^{-3}$ for MA, $p = 0.092$ for MB; two-tailed Wilcoxon signed-rank test).

The second simulation examined the effect of response variance change. In this simulation, we assumed that V4 inactivation only affected response variances, without altering mean firing rates or r_{sc} structure. Under these conditions, population d' remained largely unaffected (Fig. 5a, b, 4th column; blue vs. green bar, $p = 0.054$ for MA, $p = 0.024$ for MB).

The last simulation examined the effect of firing rate change. This simulation revealed that changes in V1 neuronal firing rates alone after V4 inactivation fully accounted for the observed decrease in population d' (Fig. 5a, b, last column; blue vs. green bar, $p = 2.3 \times 10^{-3}$ for MA, $p = 1.1 \times 10^{-4}$ for MB). These results closely match those in the first two columns.

These simulation results support the view that visual binding or grouping is primarily mediated by changes in firing rates within neuronal ensembles⁴¹.

Similar feedback modulation mode across distinct grouping and segmentation processes

Previous studies have shown that local singleton detection²⁶ and global surface segmentation^{3,42} also involve late response components in V1.

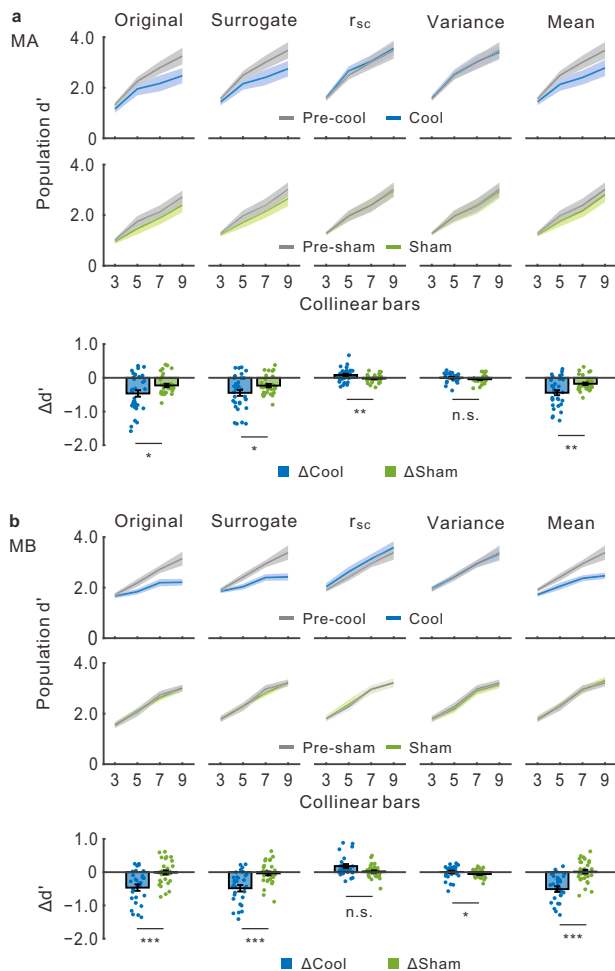


Fig. 5 | Dissecting feedback influences on V1 neural population code.

a Decoding performance (population d') based on FLD analysis of the original dataset (1st column); a surrogate dataset constructed using original key metrics (2nd column); and three simulated datasets assuming that V4 inactivation altered only the r_{sc} structure, response variances, or mean firing rates (columns 3–5). Analyses were performed for cooling experiments (top row) and sham-cooling controls (middle row) to assess the true effects of V4 inactivation (bottom row). Results for each contour length and each contour position were analyzed separately. By averaging across different contour positions ($n = 8$), the mean (solid curve) and SEM (light shading) were calculated as a function of contour length (top and middle rows). By averaging across the 4 contour lengths and 8 contour positions (bottom row), the mean changes ($\Delta d'$) due to V4 cooling ($\Delta Cool$) and sham cooling ($\Delta Sham$) were computed (blue vs. green bars, $n = 32$ conditions, 8 contour positions \times 4 contour lengths; from left to right, $p = 0.023$, $p = 0.020$, $p = 6.3 \times 10^{-3}$, $p = 0.054$, $p = 2.3 \times 10^{-3}$, two-tailed Wilcoxon signed-rank test). **b** Similar analysis to **a**, for monkey MB (upper and middle rows: $n = 8$; bottom row: $n = 32$, from left to right, $p = 9.3 \times 10^{-4}$, $p = 4.7 \times 10^{-4}$, $p = 0.092$, $p = 0.024$, $p = 1.1 \times 10^{-4}$). Error bars represent SEM; n.s., $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Source data are provided as a Source Data file.

However, it remains unclear whether the feedback modulation observed in contour grouping similarly plays a crucial role in these distinct grouping and segmentation processes. To investigate this, we performed a surgical ablation of the V4 region corresponding to the lower right visual field—opposite to that targeted in cooling experiments—in monkey MB (Supplementary Fig. S1e, f). A 10×10 micro-electrode array was implanted in the V1 region retinotopically matched to the lesioned V4 area. This approach was chosen to address potential limitations of cooling-induced silencing, which may not fully eliminate V4-to-V1 feedback originating from deep cortical layers (see Fig. 1d, third panel, the IG layer). We then compared behavioral and V1

responses in contour, singleton, and surface detection tasks before and after the lesion.

In the contour detection task, behavioral performance was severely impaired after the V4 lesion (Fig. 6a; $p = 6.0 \times 10^{-6}$, two-tailed Wilcoxon signed-rank test). The mean behavioral d' dropped by 82.9% compared to a 46.7% net reduction observed during the cooling experiment in the same animal (see Fig. 2e, f). This severe impairment was restricted to the visual-field quadrant corresponding to the lesioned V4 area (Supplementary Fig. S5). Correspondingly, contour-induced modulations of V1 late responses (90–200 ms) were markedly reduced (Fig. 6b, c, upper panel: near sites, red vs. gray, $p = 0.024$, mean reduction of 107.7%; lower panel: far sites, $p = 1.1 \times 10^{-3}$, 44.7% mean reduction; two-tailed Wilcoxon rank-sum test). Notably, the post-lesion mean d' of near sites dropped to nearly zero across all contour lengths (Fig. 6c, upper, red). While the V4 lesion had a much greater impact on contour grouping in V1 compared to the cooling approach, the estimated latencies of V4-to-V1 feedback were similar in both cases (see Supplementary Fig. S6 and Fig. 4 for comparison). The near-complete loss of contour facilitation and the substantial reduction in background inhibition further corroborate the critical role of V4 feedback in engaging V1 late responses for processing global contours.

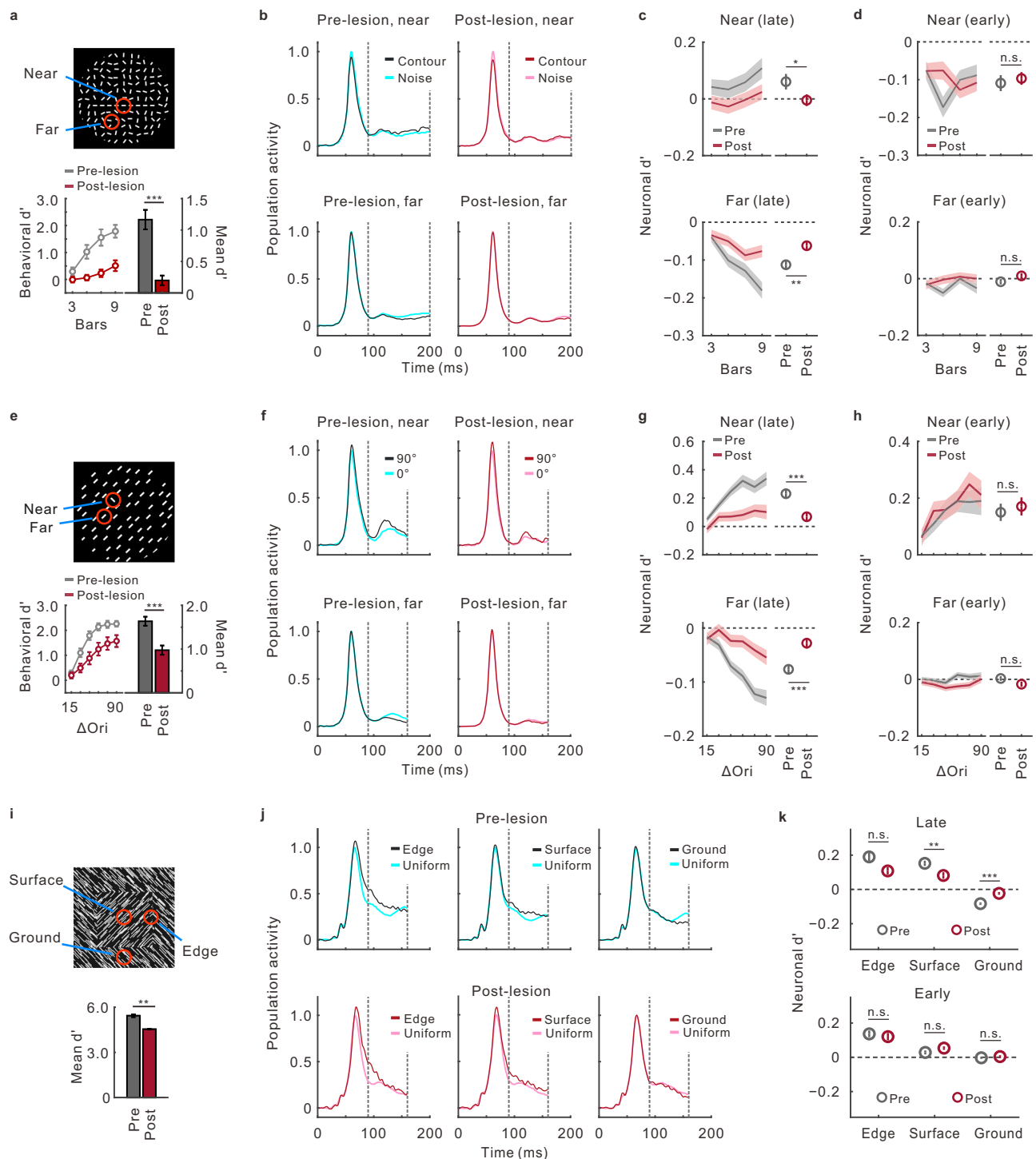
In contrast, early V1 responses (0–90 ms) were unaffected by either contour length or the V4 lesion (Fig. 6d, pre- vs. post-lesion, $p = 0.77$ and $p = 0.12$ for near and far sites, respectively; two-tailed Wilcoxon rank-sum test). Across different contour lengths, the d' values remained around 0 for far sites but consistently below 0 for near sites (Fig. 6d, lower vs. upper panel). This difference between near and far sites during the early response epoch likely reflects orientation-dependent contextual inhibition intrinsic to V1, as previously reported¹⁷ (see also Discussion).

As in the cooling experiments, the lesion experiment results were based on V1 sites pooled across eight contour positions, potentially including repeated electrode measurements. We accounted for this dependency using linear mixed-effects models, which yielded consistent results (Supplementary Table S3).

In the singleton detection task (Fig. 6e, top), the monkey indicated the presence or absence of an orientation singleton by making a saccade in Go trials or holding fixation in No-go trials (see Methods for details). The V4 lesion caused a significant decrease in behavioral d' (Fig. 6e, bottom, red vs. gray), with a mean reduction of 40.4% ($p < 10^{-6}$, two-tailed Wilcoxon signed-rank test). This impairment was about half as severe as that observed in global contour detection (compared with Fig. 6a), indicating a relatively preserved ability to perceive local feature contrasts. In fact, even in the contour detection task, when the collinear bars were highlighted by luminance contrast against the background, the animal was able to reliably detect the highlighted bars in the affected visual field, with performance largely independent of contour length (Supplementary Fig. S5). These observations are consistent with previous findings demonstrating that V4 lesions have minimal effects on detecting luminance-defined cues^{36,37} and isolated simple shapes⁴³, in contrast to the severe impairments observed in detecting cues requiring grouping and segmentation³⁷.

For V1 sites, we categorized them into near and far groups based on their RF locations relative to the singleton (Methods). It is important to note that for an individual orientation-selective neuron with its RF on the singleton, responses are influenced by two intertwined factors: the orientation contrast between the singleton and background (contextual cue) and the singleton's orientation relative to the neuron's preferred orientation (neuronal feature selectivity). However, the orientation contrast signal can be isolated from population-averaged responses^{26,44} (Methods, Isolation of orientation contrast signals in singleton detection).

Before the V4 lesion, we replicated previous results²⁶. Compared to the singleton-absent condition, the population PSTH of V1 near sites in response to the most salient singleton (orthogonal to the background) was elevated from the response outset (Fig. 6f, upper left).



Singleton facilitation increased with orientation contrast during both the late (90-160 ms, Fig. 6g, upper, gray curve) and early (0-90 ms, Fig. 6h, upper, gray) epochs. For far sites, late responses exhibited background inhibition (Fig. 6f, lower left), which increased with orientation contrast (Fig. 6g, lower, gray), while early responses were unaffected by the singleton located outside their RFs (Fig. 6h, lower, flat gray curve).

After the V4 lesion, both singleton facilitation and background inhibition in the late response epoch were significantly diminished (Fig. 6f, right vs. left; Fig. 6g, upper panel: near sites, red vs. gray, $p = 2.6 \times 10^{-4}$, mean reduction of 70.3%; lower panel: far sites, $p = 7.6 \times 10^{-5}$, mean reduction of 68.0%; two-tailed Wilcoxon rank-sum test). By contrast, the early response epoch, where singleton

facilitation had already emerged, was unaffected by the lesion (Fig. 6h, upper panel: near sites, red vs. gray, $p = 0.41$; lower panel: far sites, $p = 0.12$). Linear mixed-effects models accounting for repeated sampling from the same electrodes confirmed these results (Supplementary Table S4). This suggests that, despite the early, bottom-up origin of orientation contrast signals in V1 near sites (Fig. 6h, upper), V4 feedback is crucial for amplifying the singleton signal (Fig. 6g, upper) and suppressing background interference (Fig. 6g, lower), thereby enhancing figure-ground contrast signals in V1. This feedback modulation pattern mirrors that observed in contour grouping (compare Fig. 6e-h with 6a-d) and extends to surface segmentation.

In the surface detection task, the monkey reported the presence (in Go trials) or absence (in No-go trials) of a square surface defined by

Fig. 6 | Impact of V4 lesion on different grouping and segmentation processes. **a–d** Contour detection. **a** Behavioral d' plotted against contour length (pre- vs. post-lesion; $n = 8$ contour positions). Bar graph: d' averaged across 4 contour lengths and 8 positions ($n = 32$, $p = 6.0 \times 10^{-6}$, two-tailed Wilcoxon signed-rank test). **b** V1 population PSTHs (9-bar contour vs. noise pattern). Upper-left, $n = 147$; upper-right, $n = 178$; lower-left, $n = 271$; lower-right, $n = 314$. Dashed lines define late epoch. **c** Late response epoch. Curves: V1 neuronal d' plotted against contour length. Circles: pre- vs. post-lesion d' averaged across contour lengths (two-tailed Wilcoxon rank-sum test; near sites, $p = 0.024$; far sites, $p = 1.1 \times 10^{-3}$; n as in **b**). **d** Similar to **c**, for early epoch. Pre- vs. post-lesion: near sites, $p = 0.77$; far sites, $p = 0.12$. **e–h** Singleton detection. **e** Behavioral d' plotted against orientation contrast ΔOri (pre- vs. post-lesion, $n = 12$, pooling 2 background orientations and 6 singleton locations). Bar graph: d' further averaged across 6 ΔOri ($n = 72$ conditions, $p < 10^{-6}$, two-tailed Wilcoxon signed-rank test). **f** V1 population PSTHs (90° vs. 0° ΔOri).

Upper-left, $n = 79$; upper-right, $n = 60$; lower-left, $n = 536$; lower-right, $n = 475$. Dashed lines: late epoch. **g** Late response epoch. Curves: V1 neuronal d' plotted against ΔOri . Circles: pre- vs. post-lesion d' averaged across ΔOri (two-tailed Wilcoxon rank-sum test; near sites, $p = 2.6 \times 10^{-4}$; far sites, $p = 7.6 \times 10^{-5}$; n as in **f**). **h** Similar to **g**, for early epoch. Pre- vs. post-lesion: near sites, $p = 0.41$; far sites, $p = 0.12$. **i–k** Surface detection. **i** Behavioral performance ($n = 9$ surface locations, Methods; pre- vs. post-lesion, $p = 0.0039$, two-tailed Wilcoxon signed-rank test). **j** Population PSTHs of three site groups (inset in **i**). Uniform, surface-absent conditions. Upper-left, $n = 90$; upper-middle, $n = 42$; upper-right: $n = 284$; lower-left, $n = 73$; lower-middle, $n = 33$; lower-right, $n = 220$. Dashed lines: late epoch. **k** Pre- vs. post-lesion neuronal d' . Upper: late epoch (left to right, $p = 0.087$, $p = 0.0028$ and $p < 10^{-6}$, two-tailed Wilcoxon rank-sum test; n as in **j**); lower: early epoch ($p = 0.83$, $p = 0.13$ and $p = 0.32$). Error bars and light shading represent SEM. n.s., $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Source data are provided as a Source Data file.

oriented textures orthogonal to the background textures (Fig. 6i, top; Methods). Unlike contour and singleton detection, the V4 lesion only slightly impaired behavioral performance in this task (Fig. 6i, bottom, red vs. gray, $p = 0.0039$, mean reduction of 16.5%; two-tailed Wilcoxon signed-rank test). The largely preserved behavioral sensitivity is likely due to the square boundaries formed by local orientation contrasts, operating similarly to orientation singletons.

V1 sites were classified into edge, surface, and ground groups based on their RF locations relative to the surface (Fig. 6i, top; Methods). The neuronal d' of each site in each group was computed relative to the uniform texture background in No-go trials. Before the V4 lesion, V1 responses to surface stimuli were consistent with previous reports^{3,5,42}, showing early edge facilitation (Fig. 6j, upper left, black vs. cyan curve), followed by surface facilitation (Fig. 6j, upper middle), and then background inhibition (Fig. 6j, upper right).

To quantify the effects of V4 lesion on surface segmentation processes in V1, we divided neuronal responses into the early (0–90 ms) and late (90–160 ms) epochs. In the late epoch, V4 ablation significantly reduced surface facilitation and background inhibition, while the reduction in edge facilitation did not reach statistical significance (Fig. 6j, lower vs. upper row; Fig. 6k, upper, red vs. gray: left, edge facilitation, $p = 0.087$, 43.1% decrease; middle, surface facilitation, $p = 0.0028$, 46.6% decrease; right, background inhibition, $p = 3.8 \times 10^{-7}$, 72.2% decrease). In contrast, no significant lesion effects were observed during the early response epoch (Fig. 6k, lower; all $ps > 0.05$). This is reflected in the pre- and post-lesion d' values, which remained around 0 for V1 ground sites (Fig. 6k, lower right), as well as the negligible pre- and post-lesion difference in early facilitatory effects for edge sites (lower left) and surface sites (lower middle). These early facilitatory effects likely resulted from the 90° orientation contrast between the center and surround textures, akin to early singleton facilitation that was unaffected by V4 lesion (see Fig. 6h, upper panel). Linear mixed-effects models, which accounted for repeated sampling from the same electrodes, confirmed these results (Supplementary Table S5). These findings indicate that V4 feedback is essential for effective surface segmentation during V1 late responses, in line with a previous report in which large lesions were made in peristriate areas⁴⁵. Our results also align with previous modeling studies suggesting that early detection processes based on local feature contrasts operate independently of feedback from higher-order cortical areas, whereas late figure-ground segregation signals critically depend on feedback modulation^{3,46}.

Overall, the lesion experiments suggest that the grouping and segmentation processes, whether based on local feature contrasts or global feature integration, are influenced by a unified mode of feedback modulation: delayed facilitation for foreground features and further delayed inhibition for background features.

Two distinct forms of contextual modulation

The unified feedback mode influencing V1 late responses is crucial for visual grouping and segmentation. In contrast, the orientation contrast

signals for singleton and surface stimuli observed during V1 early responses, which are unaffected by V4 feedback, indicate a distinct form of contextual modulation. To further differentiate these two types of modulation, we examined to what extent early and late singleton signals depend on V1 neuronal orientation selectivity, an emerging property in V1 derived from feedforward processing^{47,48}.

V1 near sites recorded during the singleton detection task were sorted by their orientation selectivity index (Methods; see also Supplementary Fig. S7b) and evenly split into better- and worse-selective groups. For each group, we constructed population-averaged PSTHs to average out the influence of individual sites' orientation tuning, thereby isolating orientation contrast signals at the population level (Fig. 7a; Methods, Isolation of orientation contrast signals in singleton detection). These PSTHs were then used to compare the singleton signals in V1 during early and late response epochs (Fig. 7b), focusing on three key factors: V4 feedback (pre- vs. post-lesion), neuronal orientation selectivity (better- vs. worse-selective V1 sites), and singleton orientation contrast (0° – 90°). A linear regression model was employed to quantify the effects of these factors—Lesion, Selectivity, and Contrast—on the early and late singleton signals, respectively (for details see Methods).

Analysis of the early response epoch using a linear model based on the data shown in Fig. 7b (left column) revealed significant effects of Contrast ($\beta = 3.6 \times 10^{-4}$, $p = 1.4 \times 10^{-7}$) and its interaction with Selectivity ($\beta = -2.2 \times 10^{-4}$, $p = 0.0034$). This suggests that, while the singleton signal increases with the orientation contrast, this effect diminishes significantly for V1 sites with lower orientation selectivity (compare solid and open symbols in the upper-left panel of Fig. 7b). No significant effects were found for the other factors and interactions (all $ps > 0.05$ for the regression coefficients), indicating that Lesion and its interactions with Selectivity and Contrast had minimal impact, as shown by the close similarity between pre- and post-lesion results (Fig. 7b, upper-left vs. lower-left panel). The early emergence of the singleton signal, its dependence on neuronal orientation selectivity, and its insensitivity to the V4 lesion (i.e., V4-to-V1 feedback) suggest that the orientation contrast signal during early V1 responses involves orientation-dependent contextual modulation, likely mediated by V1 horizontal connections^{22,28,49,50}.

By contrast, a linear model analyzing the late singleton signals (Fig. 7b, right column) revealed a significant effect of Contrast ($\beta = 7.2 \times 10^{-4}$, $p = 1.9 \times 10^{-7}$); however, unlike the early signals, a significant interaction was found between Lesion and Contrast ($\beta = -5.5 \times 10^{-4}$, $p = 5.2 \times 10^{-4}$). This indicates that although both the late and early singleton signals depend on orientation contrast, the late signal is critically dependent on V4-to-V1 feedback, as evidenced by the flattened curves of signal strength as a function of orientation contrasts (Fig. 7b, right column, lower vs. upper panel). All other factors and interactions did not reach statistical significance (all $ps > 0.05$). The absence of significant effects for Selectivity and its interaction with Lesion suggests that the late singleton signal—and thus V4-to-V1

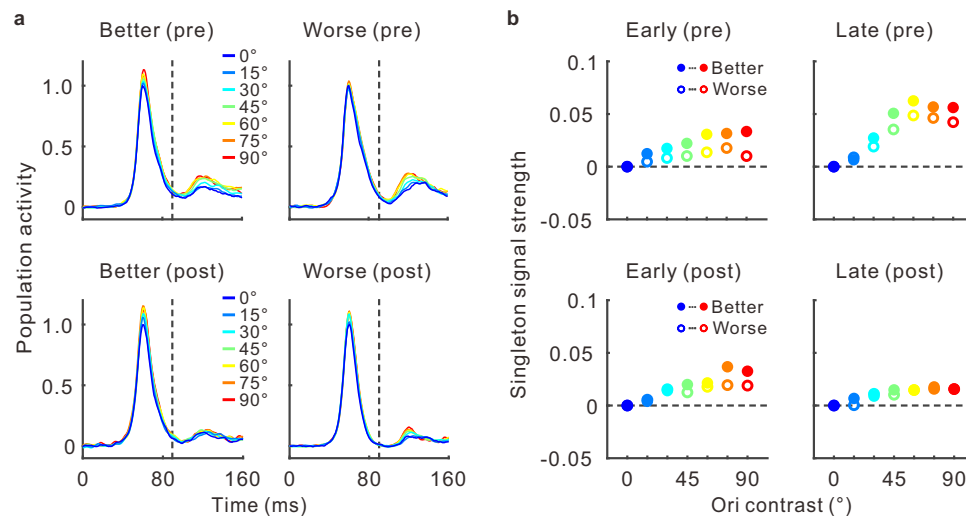


Fig. 7 | Dissociation of two forms of contextual modulation. **a** Population PSTHs for singletons of varying orientation contrasts (color-coded curves) before and after the V4 lesion. V1 near sites with RFs on the singleton were evenly divided into better and worse groups based on their orientation selectivity. Top row: pre-lesion better and worse sites ($n = 39$ in each group); bottom row: post-lesion better and worse sites ($n = 30$ in each group). Vertical dashed line at 90 ms marks the division between early and late epochs. **b** Singleton signal strength in V1 as a function of

orientation contrast, comparing early and late epochs (two columns), before and after V4 lesion (two rows), and between better and worse V1 sites (solid and open symbols). Singleton signal strength was computed by directly subtracting the PSTH for 0° contrast from the PSTHs of other contrasts. The resulting differential PSTHs for each orientation contrast were then averaged within the early or late epoch to quantify the singleton signal. The color notations are consistent with those in (a). Source data are provided as a Source Data file.

feedback modulation—is independent of neuronal orientation selectivity.

These analyses distinguish two distinct forms of contextual modulation mediated by different processes.

Discussion

Previous evidence on the role of corticocortical feedback in visual grouping and segmentation has often been speculative or inferred indirectly. In contrast, this study provides direct evidence linking V4 feedback to the ability of V1 neurons to group and segregate image elements across different stimulus paradigms featuring distinct contextual cues. Our results demonstrate that V4 feedback consistently modulates global contour grouping, surface segmentation, and local feature-contrast detection, primarily influencing the late response phases of V1 neurons. This modulation enhances the processing of foreground elements while suppressing background ones. Notably, V4-to-V1 feedback operates largely independently of V1 neurons' selectivity for local stimulus features, distinguishing it from the feature-dependent contextual modulations typically attributed to intracortical horizontal connections. This distinction highlights fundamental differences between the mechanisms underlying contextual modulation in figure-ground organization and those involved in center-surround modulation. Specifically, an initial process engages the early responses of feature-selective V1 neurons to analyze relatively local contextual cues, while a delayed feedback process engages V1 later responses to parse and integrate broader global figure-ground information, independent of local features or neuronal selectivity.

Possible mechanistic account of feedback influences on grouping and segmentation

Previous studies have proposed that higher-level visual areas integrate local elements to form a coarse global representation by pooling feedforward inputs, which then constrains subsequent feedforward processing via feedback influences. This recurrent bottom-up and top-down framework helps resolve ambiguity in complex visual scenes⁵¹. Supporting this idea, simultaneous recordings across multiple visual areas have provided insights into bidirectional inter-areal interactions during contour grouping^{4,7}.

In this study, we identified two distinct feedback modulatory effects: delayed facilitation of foreground elements and further delayed inhibition of background elements. These complementary feedback effects were dissociable in terms of spatial distribution, timing, polarity, and magnitude. Together with existing evidence, our findings offer insights into the circuit-level mechanisms underlying feedback modulation in visual grouping and segmentation.

Two potential circuit routes may explain background inhibition, either directly or indirectly. The indirect route presumes that corticocortical feedback enhances figure-ground contrast primarily by facilitating V1 neurons representing foreground elements (schematic in Supplementary Fig. S8a, left). This facilitation would activate V1 interneurons, which in turn inhibit neurons representing background stimuli near the foreground (schematic in Supplementary Fig. S8a, right). In this case, background inhibition would covary with foreground facilitation; however, the asymmetric effects of V4 cooling on these two modulatory effects (Fig. 3b, d; Supplementary Table S1) suggest otherwise. This asymmetry points to a direct-inhibition route (schematic in Supplementary Fig. S8b, right), where background inhibition is mediated via feedback connections that directly target V1 interneurons, with a modulation strength that could be independent of foreground facilitation. This is supported by studies showing that feedback projections to mouse V1 target both interneurons and pyramidal cells^{15,52,53}. Therefore, our results, combined with existing evidence, argue against the indirect route and instead support the direct feedback route as the underlying mechanism for the background inhibition observed in V1.

For foreground facilitation, two circuit routes are plausible: one directly facilitates pyramidal cells representing the foreground (schematic in Supplementary Fig. S8b, left), while the other involves a disinhibitory circuit formed by vasoactive intestinal peptide-expressing (VIP) and somatostatin-expressing (SST) interneurons in V1 (schematic in Supplementary Fig. S8c, left)^{15,54}. Recent optogenetic studies in mice support the role of the disinhibitory route in mediating center-surround feature contrast^{18,55}. Although both routes may contribute to delayed foreground facilitation, the disinhibitory route alone—without direct facilitation—would contradict our observation that foreground facilitation precedes background inhibition: Disinhibition involving

two interneurons would lead to a facilitatory effect occurring later than the inhibitory effect produced by the direct-inhibition route, which involves only one interneuron. This suggests that the direct-facilitation route is necessary, with the disinhibitory route playing a supplementary role, likely becoming active only after the onset of background inhibition.

V4 lesion experiments further underscore the crucial role of V4 feedback in visual grouping and segmentation. However, some contextual modulations in V1 persisted after the V4 lesion (Fig. 6). These residual effects likely originate from other sources, such as horizontal connections within V1 and feedback from V2 and dorsal visual areas. Nonetheless, the near-complete loss of delayed contour and surface signals following V4 lesion suggests that these other sources play a minimal role in grouping and segmenting static global features. The remaining contextual modulations in V1 are likely driven by horizontal connections, which are well-suited for mediating feature-dependent, feedforward local feature contrasts, such as orientation-contrast cues in singletons and textual boundaries. This form of modulation has been shown to persist despite V2 inactivation in anesthetized macaques⁵⁶. While contributions from visual areas higher than V4, such as the inferior temporal cortex, cannot be excluded, our findings underscore the pivotal role of V4 feedback in visual grouping and segmentation.

Feedback signals dedicated to visual grouping and segmentation

Basic response properties of V1 neurons, such as RF sizes and orientation selectivity, were minimally affected by the V4 lesion (Supplementary Fig. S7). Beyond basic visual features, V1 neurons also encode higher-order information (for reviews, see refs. 57,58), including attention, expectation, working memory, specific tasks, and stimulus context—the focus of the current study. We argue that the feedback influences on visual grouping and segmentation are distinct from top-down attentional modulation mediated by frontoparietal cortices and surround modulation mediated by feedback and horizontal connections.

Attentional control by the frontoparietal network is essential for various perceptual tasks^{59–63}, including contour and surface perception. Both contour¹ and surface³ signals in V1 are substantially attenuated under unattended conditions. This explains the observations in a recent two-photon imaging study, which demonstrated rather weak surface signals in V1 of fixating monkeys⁶⁴. While top-down attention can increase or decrease the activity of V1 neurons^{65–67} and produce a center-surround profile resembling foreground facilitation and background inhibition^{10,68,69}, the feedback signals in visual grouping and segmentation are not merely manifestations of top-down attentional control. Figure-ground contrast signals, delayed by ~50 ms relative to stimulus-evoked responses, precede attentional signals seen in V1 by ~100 ms for surface³ and contour⁷ stimuli. This temporal dissociation suggests that feedback signals for attentional selection and those for visual grouping and segmentation are fundamentally different. Furthermore, perceptual learning significantly enhances the contour^{38,70} and orientation singleton²⁶ signals in V1 late responses, while training to detect a target among distractors usually reduces frontoparietal activity and automates the task⁷¹, further suggesting a dissociation between figure-ground modulation and attentional influence.

The foreground facilitation and background inhibition observed in V1 late responses are also distinct from surround modulations—a range of contextual interactions where neuronal responses to a stimulus in the RF are modulated by surrounding stimuli. Surround modulation is likely mediated by both feedback and horizontal connections (for a review see ref. 23), but it fundamentally differs from feedback modulation in visual grouping and segmentation. Surround modulation starts early, within 0–30 ms relative to stimulus-evoked responses²³, whereas V4-to-V1 feedback

modulations occur significantly later. Additionally, surround modulation is usually feature-dependent, meaning that its strength and polarity are influenced by neuronal feature selectivity and the similarity or dissimilarity in basic visual features between stimuli inside and outside the RF; however, the feedback modulations observed in V1 late responses seem to lack such feature dependence (Fig. 7; also suggested in refs. 4,72).

Surround suppression, or iso-feature inhibition, is a prominent example of feature-dependent contextual modulation. V1 neurons' responses to a stimulus in the RF are suppressed by surrounding stimuli, with the inhibition strength proportional to the feature similarity between center and surround stimuli. Although feedback from V2^{11,12}, V3¹¹, and MT⁹ contributes to surround suppression in V1, it is distinct from the background inhibition observed during visual grouping and segmentation. For example, in processing of the orientation singleton (Fig. 6e inset), increasing the orientation contrast would decrease iso-orientation inhibition not only for near sites but also for the far sites adjacent to the singleton, opposite to increased background inhibition of far sites observed during V1 late responses. While surround suppression does not contribute to delayed background inhibition, it is closely related to the early emergence of singleton facilitation, because iso-orientation inhibition is the weakest at the singleton location; as a result, the singleton elicits relatively stronger responses for V1 near sites compared to far sites on iso-orientated background bars. In a similar vein, in processing of the texture surface defined by orientation contrast (Fig. 6i, inset), early facilitatory effects can be observed for V1 sites on the foreground surface or edges but not for those in the background (Fig. 6j, k). In fact, surround suppression is also present during early responses of V1 near sites to the collinear contours, compared to responses to the noise pattern. This is consistently shown as a small difference between the peaks of paired PSTHs (Fig. 4a, e; Fig. 6b, upper row; see also Supplementary Fig. S2a, e; Fig. S6, left column) and indicated by negative d' values (Fig. 6d, upper panel).

A limitation of this study is that the lesion experiments were conducted in only one monkey, raising potential concerns about the reproducibility of the findings, particularly regarding the consistency observed across the three figure-ground perceptual tasks. However, previous studies on intact monkeys have consistently shown similarly delayed V1 responses to the three types of figure-ground stimuli defined by distinct cues. In our study, the pre-lesion results from the single animal closely align with these established findings on contour grouping, singleton detection, and surface segmentation. Most importantly, the post-lesion results from these distinct grouping and segmentation tasks exhibit similar patterns, suggesting a shared feedback modulatory mechanism. Although the lesion experiments were limited to one subject, our findings, when combined and considered alongside existing evidence, suggest that foreground facilitation and background inhibition are likely driven by distinct feedback modulations that consistently operate across various grouping and segmentation processes. These results provide insights into counter-current processing and lay a foundation for future investigations into the circuit-level mechanisms underlying feedback modulation in visual scene analysis.

Methods

Animal preparations

Two adult male monkeys (*Macaca Mulatta*, 7–8.5 kg) participated in the experiments. All procedures complied with the National Institutes of Health Guide for Care and Use of Laboratory Animals. Ethical approval was granted by the Animal Care and Use Committee of Beijing Normal University.

The monkeys underwent aseptic surgeries under anesthesia (1.0%–2.0% isoflurane mixed in O₂). Prior to behavioral training, a titanium headpost was affixed to the skull to restrain head movement during the experiments.

V4 reversible inactivation experiments. For the reversible inactivation experiments, a craniotomy was performed over V4 in the right hemisphere, and a circle titanium chamber was secured over the craniotomy using bone cement (PALACOS R + G). The exposed dura was removed and replaced with a silicone artificial dura. A hydraulic cooling system, adapted from a previous study³⁴, was constructed, with the coolant circulation controlled automatically by an Arduino microcontroller (<https://www.arduino.cc>). To lower cortical temperature, a small cryo-loop made of stainless-steel tubing (0.7 mm outer diameter) was positioned beneath the artificial dura, shaped to fit the V4 region between the lunate and superior temporal sulcus (Fig. 1b; Supplementary Fig. S1a, c). The coolant, ethanol, circulated through the cooling system and passed through a dry ice bath. The temperature of the cryo-loop was maintained at 2.2 ± 0.3 °C, effectively silencing neuronal spiking activity³⁴, by adjusting the flow rate and monitoring it with an attached thermocouple. To assess the effects of V4 cooling on V1 responses, a microelectrode array (10 × 10 electrodes, 0.5 mm electrode length, 0.4 mm inter-electrode spacing; Blackrock Neurotech, USA) was implanted in V1 region retinotopically matched the manipulated V4 region (Supplementary Fig. S1b, d).

V4 lesion experiments. For the V4 lesion experiments on monkey MB, we targeted the left brain hemisphere, which was opposite to that tested in the cooling experiments. An electrode array—with the same specifications as those used in the cooling experiments—was implanted in the V1 region retinotopically matched the lesioned V4 area (Supplementary Fig. S1e, f). V4 lesions were made by suction aspiration of the gray matter, extending from the tip of the inferior occipital sulcus to approximately 10 mm dorsally, between the lunate sulcus and the superior temporal sulcus.

Electrophysiological recordings

Neuronal responses were recorded using a 128-channel Cerebus System (Blackrock Neurotech, USA). Raw electrophysiology signals were band-pass filtered between 250 and 7500 Hz and sampled at 30 kHz. Multi-unit activity was detected using a root-mean-square (RMS) voltage threshold with a signal-to-noise ratio of 4.0.

Receptive field mapping. The receptive fields (RFs) of V1 recording sites were mapped using square-wave gratings (2 cycles per degree, drifting at 3 cycles per second). The gratings, presented within a narrow band ($0.3^\circ \times 6^\circ$ vertically or $6^\circ \times 0.3^\circ$ horizontally), were moved horizontally or vertically across the screen to determine the RF center and size for each V1 site. A Gaussian function was fitted to the position tuning curve of each site, with RF size defined as 2×1.96 SD of the Gaussian fit. Only sites with reliable RF profiles (Gaussian fit $R^2 > 0.7$) for both horizontal and vertical position tuning were chosen as valid sites and included in the dataset. The number of valid V1 sites fluctuated across experiments conducted on different days. In the cooling and sham-cooling experiments, this number ranged from 78 to 92 for MA and from 76 to 79 for MB. In the lesion experiments, the number varied from 41 to 71 for MB across the pre- and post-lesion sessions. For a valid V1 site, the RF sizes along the horizontal and vertical dimensions were averaged to quantify the RF size of the site. In the cooling experiments conducted on 8 days, we performed RF mapping before each pre-cooling session. If a V1 site was identified as valid on multiple days, its mean RF size was calculated across days. Across all valid sites, the mean RF size (mean \pm SD) was $1.69^\circ \pm 0.39^\circ$ for monkey MA ($n = 94$) and $1.16^\circ \pm 0.39^\circ$ for MB ($n = 80$). In the lesion experiments on MB, we mapped the pre- and post-lesion RFs and counted the sites that were valid both before and after the lesion. The mean pre- and post-lesion RF sizes were $1.24^\circ \pm 0.29^\circ$ and $1.30^\circ \pm 0.25^\circ$, respectively ($n = 82$; Supplementary Fig. S7a).

Using a similar RF mapping approach and before conducting any experiments, a few single-electrode penetrations were made in V4 to

verify the areas targeted by cooling and lesions (Supplementary Fig. S1).

To assess the orientation selectivity of valid V1 sites (as defined above), gratings with varying orientations (in 22.5° increments) were presented within a 4° diameter circular aperture centered on the cluster of V1 RFs. This generated an orientation tuning curve for each V1 site, from which the preferred orientation was defined as the orientation eliciting the maximum response. Orientation selectivity was quantified using the orthogonal-to-preferred (O/P) ratio, calculated as the response to the orthogonal orientation divided by the response to the preferred orientation after subtracting the mean spontaneous activity⁷³. The O/P ratio ranges from 0 to 1, with lower values indicating stronger orientation selectivity. In the cooling experiments, each V1 site's O/P ratio was averaged across all pre-cooling sessions. Across all sites, the mean O/P ratio (mean \pm SD) was 0.45 ± 0.22 for monkey MA ($n = 94$) and 0.52 ± 0.22 for MB ($n = 80$). In the lesion experiments on MB, the mean pre-lesion and post-lesion O/P ratios were 0.63 ± 0.22 and 0.64 ± 0.19 , respectively ($n = 82$; Supplementary Fig. S7b).

Validation of V4 reversible inactivation. To validate the effectiveness and reversibility of the cooling setup, we used a 24-channel linear electrode array (U-Probe, 100 μ m inter-electrode spacing, Plexon Inc.), inserted perpendicularly into area V4 with a microelectrode drive (NAN Instruments, Israel). The relative depth of the electrode recording sites within the cortex was estimated using current source density (CSD) analysis of local field potentials (LFPs), which were low-pass filtered at 250 Hz and sampled at 2 kHz. LFPs were evoked by presenting a full-screen checkerboard pattern, and the CSD was computed using the formula⁷⁴:

$$\text{CSD}(x) = -\frac{\phi(x-h) - 2\phi(x) + \phi(x+h)}{h^2} \quad (1)$$

where x represents the cortical depth of a recording site, ϕ is the potential, and h is the spacing between adjacent sites. The CSD was calculated at 1-ms intervals after stimulus onset. The V4 input layer (i.e., the granular layer, G) was identified based on the spatiotemporal sink-source profile⁷⁵, corresponding to the cortical depth range of the earliest current sink⁷⁶. The supra-granular layer (SG) and infra-granular layer (IG) were subsequently determined (Fig. 1d, left).

Behavioral tasks and visual stimuli

In the V4 cooling experiments, both monkeys performed a contour detection task in the lower left visual field (Fig. 1c). In a separate set of experiments involving the V4 lesion in monkey MB, three figure-ground perceptual tasks—contour, singleton, and surface detection—were compared in the lower right visual field. In all tasks, the visual stimuli were generated by a stimulus generator (ViSaGe MKII; Cambridge Research System) and displayed on a gamma-corrected CRT monitor (Iiyama Vision Master Pro 514, 1200 × 900 pixels at 100 Hz). The monkeys were seated 100 cm from the monitor, with eye positions sampled at 500 Hz by an infrared tracking system⁷⁷. Each trial began with the monkey fixating within an invisible 1.2° diameter window around the fixation point. After the monkey kept fixation for 1–1.2 seconds, a stimulus pattern was presented in the visual-field area corresponding to the RFs of recorded V1 sites. A Go/No-go procedure under time pressure was used for all tasks.

During the contour detection task in the cooling experiment (Fig. 2a), a circular stimulus pattern with a diameter of 6.5° (for MA, centered at -4.4° , -4.5°) or 5.2° (for MB, centered at -2.2° , -1.4°) was displayed on a gray background (4.4 cd/m²). The pattern consisted of randomly oriented bars ($0.25^\circ \times 0.05^\circ$ or MA, $0.2^\circ \times 0.04^\circ$ for MB, 13.3 cd/m²) evenly distributed in an invisible square grid (compartment size $0.5^\circ \times 0.5^\circ$ for MA, $0.4^\circ \times 0.4^\circ$ for MB). Each bar had a random

position jitter of 0–30% of the compartment size. In Go trials, a global contour was formed within the stimulus pattern by collinearly arranging 3, 5, 7, or 9 bars (Fig. 2a, upper right inset, referred to as the contour pattern). In No-go trials, the embedded contour consisted of only one bar, making it indistinguishable from the background bars (Fig. 2a, lower right inset, referred to as the noise pattern). The contour path was off-centered by one row. The contour orientation was set by rotating the stimulus around the pattern center in multiples of 45°, generating a total of 8 possible contour positions and orientations. The monkey was trained to make a saccade to the contour pattern within 500 ms after its onset or to maintain fixation during the noise pattern presentation until it disappeared 500 ms later. A correct Go or No-go response was rewarded with a drop of liquid.

In the V4 lesion experiments, monkey MB performed three tasks—contour, singleton and surface detection—in different sessions. The contour detection task was similar to that in the V4 cooling experiments, except for the stimulus pattern's diameter (6.5°) and center location (4.4°, −4.4°). The pattern consisted of 0.25° × 0.05° bars arranged in a 0.5° grid.

In the singleton detection task (Fig. 6e), the stimuli had a diameter of 6.0° and consisted of 0.25° × 0.05° bars arranged in 0.5° compartments. In No-go trials, all bars were oriented uniformly, while in Go trials, the orientation of a designated bar was varied, creating an orientation singleton. Except for the singleton, each background bar was randomly jittered from its compartment center by 0 to 0.18°. A total of 72 stimulus conditions were tested in Go trials: 2 orientations for the iso-orientated background bar (45°, 135°); 6 orientation contrasts for the singleton (ranging from 15° to 90° in 15° increments relative to the background bars, with clockwise and counterclockwise deviations counterbalanced); and 6 singleton locations (distributed within the cluster of recorded RFs). The monkey was rewarded for making a saccade into a 2° circular window around the singleton in Go trials or maintaining fixation in No-go trials.

In the surface detection task (Fig. 6i), the textures consisted of iso-oriented line segments³, each measuring 0.5° × 0.019°. In Go-trials, a 3° × 3° square surface was embedded in a full screen of orthogonal texture, with the center-surround texture orientations counterbalanced across trials (45°/135° as illustrated in Fig. 6i, or 135°/45° by simultaneously flipping the center and surround textures). In corresponding No-go trials, only the uniform texture background was shown. A total of 9 surface locations were tested as follows: the centroid of the cluster of V1 RFs was placed at the square center, at the midpoint of the 4 edges, or outside the square in such way that the entire RF cluster was close to one of the 4 edges. The monkey received a reward for making a saccade to the central 2° region of the square surface in Go trials or maintaining fixation in No-go trials.

Data analyses

Behavioral d'. Behavioral d' was used to quantify the animals' detection performance based on signal detection theory, using the following formula:

$$d' = \text{norminv}(\text{Hit rate}) - \text{norminv}(\text{False alarm rate}) \quad (2)$$

where the function *norminv* returns the inverse of the standard normal cumulative distribution function for the given hit or false alarm rate. If the hit or false alarm rate equals 0 or 1, the *norminv* function returns infinity. In such extreme cases, a correction was applied as follows⁷⁸: 0.5 was added to both the number of hit trials and false alarm trials, and 1.0 was added to both the total number of Go trials and No-go trials.

Neuronal d'. To quantify the sensitivity of a V1 site in differentiating between Go (target-present) and No-go (target-absent) stimuli, we

computed the neuronal d':

$$d' = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}} \quad (3)$$

where μ_1 , μ_2 represent the mean firing rates in the Go and No-go trials, respectively; σ_1 , σ_2 represent the corresponding standard deviations.

Effect of V4 cooling or lesion on d'. To estimate the net cooling effect, i.e., the percent change in behavioral d' or neuronal d' entirely attributable to V4 cooling, we controlled for the sham-cooling effect using the following formula:

$$\text{Net cooling effect} = \left(\frac{d'_{\text{Pre-cool}} - d'_{\text{Cool}}}{d'_{\text{Pre-cool}}} - \frac{d'_{\text{Pre-sham}} - d'_{\text{Sham}}}{d'_{\text{Pre-sham}}} \right) \times 100\% \quad (4)$$

The percent change in behavioral d' or neuronal d' due to the V4 lesion was computed as follows:

$$\text{Lesion effect} = \left(\frac{d'_{\text{Pre-lesion}} - d'_{\text{Post-lesion}}}{d'_{\text{Pre-lesion}}} \right) \times 100\% \quad (5)$$

Analysis of V1 responses. A post-stimulus time histogram (PSTH) was constructed by binning spike times in 1-ms intervals and averaging across trials for each V1 site under each stimulus condition. After subtracting the mean spontaneous activity recorded before stimulus onset (−200 to 0 ms), the raw PSTH was smoothed using a 9-ms boxcar filter. A population PSTH was then created by averaging across the same group of sites (near or far) for each stimulus condition (e.g., Supplementary Fig. S6a). To facilitate comparison, the PSTHs for different stimulus conditions were normalized so that the peak of the PSTH in target-absent (No-go) trials was set to unity (e.g., Supplementary Fig. S6b).

In the Go/No-go tasks, the monkey could make a saccadic response, terminating the trials at various time points after stimulus onset. A valid saccadic reaction time was defined as the time point within 100–500 ms after stimulus onset when the monkey's gaze deviated by 1° from the fixation point. As the probability of saccades increased over time, the actual number of trials contributing to the time bins of the PSTHs decreased accordingly. We selected a cut-off time of 200 ms for the contour stimuli and 160 ms for the singleton and surface stimuli, aiming to balance the number of contributing trials around the cut-off time and the duration of late response epochs of V1 sites. We confirmed that, at the chosen cut-off times, a sufficient number of trials remained even for the most salient stimuli (e.g., the longest 9-bar contour, Supplementary Fig. S3), which typically elicited the shortest reaction times. In our data analyses, the number of trials in the last bin among all PSTHs was at least 6.

Neuronal response latency estimation. In the contour detection task, we estimated the time point when the mean neural response to the contour pattern diverged from that to the noise pattern. This time point was defined as the latency of contour-induced modulation. Specifically, we computed the response differences within each 1-ms bin between the two population PSTHs in response to the contour pattern and the noise pattern, generating a differential PSTH (contour − noise for contour facilitation; noise − contour for background inhibition). The differential PSTH $f(t)$ was fitted using the formula³⁹:

$$f(t) = d^* \exp(\mu\alpha + 0.5\sigma^2\alpha^2 - \alpha t) * G(t, \mu + \sigma^2\alpha, \sigma) + c^* G(t, \mu, \sigma) \quad (6)$$

where $G(t, \mu, \sigma)$ is a cumulative Gaussian. The latency of contour-induced modulation was taken as the time point when the fitted curve reached 33% of its maximum. The mean and SEM of the latency were estimated by resampling the same group of V1 sites (near or far) with

replacement and repeating the above curve-fitting procedure 100 times.

Using this approach, we estimated the latencies of contour facilitation (for V1 near sites; Fig. 4a, e for 9-bar contour; Supplementary Fig. S2a, e for 5-bar contour) and background inhibition (for V1 far sites; Fig. 4b, f for 9-bar contour; Supplementary Fig. S2b, f for 5-bar contour) in V4 cooling and sham-cooling experiments. The latency of V4 feedback modulation was estimated similarly by computing a differential PSTH in which the net cooling effect was isolated by controlling for the sham-cooling effect (Fig. 4c, d, g, h; Supplementary Fig. S2, corresponding panels).

Noise correlation. For trials with the same contour length during the same recording session in the cooling and sham-cooling experiments, trial-by-trial spike counts (90–200 ms after stimulus onset) from each V1 site were z-scored by subtracting the mean response across these trials and dividing by the standard deviation. This process isolated trial-to-trial variability while minimizing the influence of firing rates. To prevent artificial correlations due to outliers, data points with z-scores greater than 3.0 were excluded^{79,80}. Subsequently, the standardized responses were further z-scored in blocks of 20 trials to minimize the confounding effects of slow variations in neural responses. Mean pairwise noise correlations were then computed between paired near-near, far-far, and near-far sites, respectively (Supplementary Fig. S4), as the Pearson correlation coefficients of the z-scored trial responses. Notably, significant cooling effects on noise correlation were observed only in contour-pattern trials, but not in noise-pattern trials that were randomly interleaved within the same recording session, further excluding the potential confounding factor of slow variations in neural responses.

Analysis of neural population code. By assigning a unique dimension to each V1 site, the collective firing activity of all sites in a trial could be represented as a vector within a high-dimensional neural population space. The Fisher's Linear Discriminant (FLD) analysis⁸¹ was used to decode, or classify, V1 population responses in Go and No-go trials for a given contour length. The FLD analysis identified an axis (the FLD axis) that maximized the separation (d') of projected population response vectors between Go and No-go trials. Before decoding, we preprocessed the population responses using principal component analysis (PCA) and kept the first 20 components. FLD analysis was then applied to the dimensionality-reduced dataset to derive the population d' , which served as a measure of the overall strength of contour signals contained in the population activity. A 10-fold cross-validation procedure was repeated 100 times to achieve a reliable estimate of decoding accuracy (mean $d' \pm \text{SEM}$).

In the simulation analyses of factors influencing population activity (Fig. 5), we modeled each stimulus condition using a multivariate normal distribution (via the Matlab function *mvrnd*), specifying the mean responses of individual V1 sites and their covariance matrix. The covariance matrix was constructed by multiplying the Pearson correlation coefficients of each pair of V1 sites by the variances of the corresponding sites. Trial-by-trial responses of all V1 sites were simulated by drawing from the multivariate normal distribution to generate surrogate datasets, which were then analyzed using the FLD method described above.

Classification of V1 sites in singleton and surface detection. In the singleton detection task, a V1 site was classified as a near site if the singleton's center fell within ± 1 SD of the Gaussian-fitted RF, or as a far site if the singleton's center was outside the RF (RF size was defined as 2×1.96 SD of the Gaussian fit, see Receptive field mapping).

In the surface detection task, 9 surface locations were tested with the RF cluster's centroid (a) located at the square's center, (b) centered on each of the 4 edges, or (c) positioned so that the entire RF cluster was outside the square adjacent to of the 4 the edges. Based on these

conditions, a V1 site was classified as a surface, edge, or ground site as follows. In condition (a), a site was defined as a surface site if the RF was centered within the central 1° of the square surface and did not overlap with any edges. In condition (b), a site was classified as an edge site if the surface edge fell within ± 1 SD of the Gaussian-fitted RF. All sites in condition (c) were categorized as ground sites.

Isolation of orientation contrast signals in singleton detection. In the singleton detection task, for each orientation contrast (θ , 0 – 90°), we combined the two background bar orientations (45° and 135°) with four possible singleton orientations ($45^\circ \pm \theta$ and $135^\circ \pm \theta$). For each V1 site, we first averaged its responses across these four conditions to obtain an individual mean, and then computed the population average by averaging these individual means across all V1 sites. When θ is 0° or 90° , the singleton and background orientations are fully balanced across the four conditions; thus, the individual mean effectively removes the influence of each site's orientation tuning and isolates the orientation contrast signal. For other θ values, the singleton and background orientations are not precisely balanced. Nevertheless, because V1 sites are tuned to a range of orientations, the population average largely cancels out orientation tuning effects, allowing the orientation contrast signal to be isolated at the population level, as demonstrated in previous studies^{26,44}.

Statistical analyses. Unless otherwise stated, standard statistical analyses were applied as appropriate, including the Wilcoxon signed-rank test for paired data and the Wilcoxon rank-sum test for unpaired data. When $p \geq 10^{-6}$, the exact p-value was reported; otherwise, $p < 10^{-6}$ was used for simplicity.

To disentangle the individual and combined effects of key factors on V1 orientation singleton signals, we employed a linear regression model (using the Matlab function *fitlm*).

In this analysis, we selected V1 sites with RFs covering the singleton and pooled data from the six tested singleton locations. These recording sites were evenly divided into two groups based on their orientation selectivity index (O/P ratio, see Receptive field mapping). For the pre-lesion experiments, responses from each group of sites were averaged separately for each orientation contrast, generating two sets of population PSTHs corresponding to the better- and worse-selective sites in response to different orientation contrasts (Fig. 7a, upper two panels). The post-lesion data were processed in the same manner (Fig. 7a, lower two panels). The population-averaged PSTHs largely averaged out the orientation tuning signals of individual V1 neurons, thereby isolating the orientation contrast signals (see Isolation of orientation contrast signals in singleton detection). From these PSTHs, we calculated the mean early (0–90 ms) and late (90–160 ms) orientation singleton signals for each orientation contrast. These signals were obtained by directly subtracting the PSTH at 0° contrast from the PSTH at each contrast. Using the mean singleton signals as the dependent variable and their associated variables as predictors, we constructed separate linear models for the early and late responses:

$$\begin{aligned} \text{Signal} \sim & \beta_0 + \beta_1 \text{Contrast} + \beta_2 \text{Selectivity} + \beta_3 \text{Lesion} \\ & + \beta_4 (\text{Contrast} \cdot \text{Selectivity}) + \beta_5 (\text{Contrast} \cdot \text{Lesion}) \\ & + \beta_6 (\text{Selectivity} \cdot \text{Lesion}) + \beta_7 (\text{Contrast} \cdot \text{Selectivity} \cdot \text{Lesion}) \end{aligned}$$

Here, Contrast, Selectivity, and Lesion were the three predictors: the singleton orientation contrast (0° to 90° in 15° steps, treated as a continuous variable); neuronal orientation selectivity (better or worse, as a binary variable); and the feedback state (pre- or post-lesion, as a binary variable). Because this analysis relied on population-averaged responses to isolate orientation contrast signals, repeated sampling by individual electrodes was not treated as a random factor. Thus, each regression model included 28 samples (7 Contrast levels $\times 2$ Selectivity $\times 2$ Lesion states).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw neural spiking datasets (stored in complex, self-customized binary formats) cannot be directly deposited in public repositories but are available from the corresponding authors upon request. Source data are provided with this paper.

Code availability

The present study used standard Matlab toolboxes for all data analyses; no custom code central to the conclusions was developed.

References

- Li, W., Piëch, V. & Gilbert, C. D. Contour saliency in primary visual cortex. *Neuron* **50**, 951–962 (2006).
- Zipser, K., Lamme, V. A. & Schiller, P. H. Contextual modulation in primary visual cortex. *J. Neurosci.* **16**, 7376–7389 (1996).
- Poort, J. et al. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* **75**, 143–156 (2012).
- Chen, M. et al. Incremental integration of global contours through interplay between visual cortical areas. *Neuron* **82**, 682–694 (2014).
- Poort, J., Self, M. W., van Vugt, B., Malkki, H. & Roelfsema, P. R. Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex. *Cereb. Cortex* **26**, 3964–3976 (2016).
- Klink, P. C., Dagnino, B., Gariel-Mathis, M. A. & Roelfsema, P. R. Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. *Neuron* **95**, 209–220 (2017).
- Chen, R., Wang, F., Liang, H. & Li, W. Synergistic processing of visual contours across cortical layers in V1 and V2. *Neuron* **96**, 1388–1402 (2017).
- Sandell, J. H. & Schiller, P. H. Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *J. Neurophysiol.* **48**, 38–48 (1982).
- Hupe, J. M. et al. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* **394**, 784–787 (1998).
- Ekstrom, L. B., Roelfsema, P. R., Arsenault, J. T., Bonmassar, G. & Vanduffel, W. Bottom-up dependent gating of frontal signals in early visual cortex. *Science* **321**, 414–417 (2008).
- Nassi, J. J., Lomber, S. G. & Born, R. T. Corticocortical feedback contributes to surround suppression in V1 of the alert primate. *J. Neurosci.* **33**, 8504–8517 (2013).
- Nurminen, L., Merlin, S., Bijanzadeh, M., Federer, F. & Angelucci, A. Top-down feedback controls spatial summation and response amplitude in primate visual cortex. *Nat. Commun.* **9**, 2281 (2018).
- Roberts, M. J. et al. Acetylcholine dynamically controls spatial integration in marmoset primary visual cortex. *J. Neurophysiol.* **93**, 2062–2072 (2005).
- Debes, S. R. & Dragoi, V. Suppressing feedback signals to visual cortex abolishes attentional modulation. *Science* **379**, 468–473 (2023).
- Zhang, S. et al. Long-range and local circuits for top-down modulation of visual cortex processing. *Science* **345**, 660–665 (2014).
- Huh, C. Y. L., Peach, J. P., Bennett, C., Vega, R. M. & Hestrin, S. Feature-specific organization of feedback pathways in mouse visual cortex. *Curr. Biol.* **28**, 114–120 (2018).
- Pak, A., Ryu, E., Li, C. & Chubukin, A. A. Top-down feedback controls the cortical representation of illusory contours in mouse primary visual cortex. *J. Neurosci.* **40**, 648–660 (2020).
- Kirchberger, L. et al. The essential role of recurrent processing for figure-ground perception in mice. *Sci. Adv.* **7**, eabe1833 (2021).
- Fisek, M. et al. Cortico-cortical feedback engages active dendrites in visual cortex. *Nature* **617**, 769–776 (2023).
- Rockland, K. S., Lund, J. S. & Humphrey, A. L. Anatomical binding of intrinsic connections in striate cortex of tree shrews (*tupaia glis*). *J. Comp. Neurol.* **209**, 41–58 (1982).
- Gilbert, C. D. & Wiesel, T. N. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133 (1983).
- Stettler, D. D., Das, A., Bennett, J. & Gilbert, C. D. Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* **36**, 739–750 (2002).
- Angelucci, A. et al. Circuits and mechanisms for surround modulation in visual cortex. *Annu. Rev. Neurosci.* **40**, 425–451 (2017).
- Knierim, J. J. & Van Essen, D. C. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* **67**, 961–980 (1992).
- Li, W., Thier, P. & Wehrhahn, C. Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *J. Neurophysiol.* **83**, 941–954 (2000).
- Yan, Y., Zhaoping, L. & Li, W. Bottom-up saliency and top-down learning in the primary visual cortex of monkeys. *Proc. Natl. Acad. Sci. USA* **115**, 10499–10504 (2018).
- Nothdurft, H. C., Gallant, J. L. & Van Essen, D. C. Response modulation by texture surround in primate area V1: Correlates of “pop-out” under anesthesia. *Vis. Neurosci.* **16**, 15–34 (1999).
- Zhaoping, L. A new framework for understanding vision from the perspective of the primary visual cortex. *Curr. Opin. Neurobiol.* **58**, 1–10 (2019).
- Donk, M. & van Zoest, W. Effects of salience are short-lived. *Psychol. Sci.* **19**, 733–739 (2008).
- Pasupathy, A. & Connor, C. E. Population coding of shape in area V4. *Nat. Neurosci.* **5**, 1332–1338 (2002).
- Rockland, K. S., Saleem, K. S. & Tanaka, K. Divergent feedback connections from areas V4 and TEO in the macaque. *Vis. Neurosci.* **11**, 579–600 (1994).
- Rockland, K. S. & Vanhoesen, G. W. Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cereb. Cortex* **4**, 300–313 (1994).
- Ungerleider, L. G., Galkin, T. W., Desimone, R. & Gattass, R. Cortical connections of area V4 in the macaque. *Cereb. Cortex* **18**, 477–499 (2008).
- Lomber, S. G., Payne, B. R. & Horel, J. A. The cryoloop: an adaptable reversible cooling deactivation method for behavioral or electrophysiological assessment of neural function. *J. Neurosci. Methods* **86**, 179–194 (1999).
- Merigan, W. H. Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. *Vis. Neurosci.* **13**, 51–60 (1996).
- De Weerd, P., Desimone, R. & Ungerleider, L. G. Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Vis. Neurosci.* **13**, 529–538 (1996).
- Merigan, W. H. Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Vis. Neurosci.* **17**, 949–958 (2000).
- Yan, Y. et al. Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nat. Neurosci.* **17**, 1380–1387 (2014).
- Khayat, P. S., Pooresmaeili, A. & Roelfsema, P. R. Time course of attentional modulation in the frontal eye field during curve tracing. *J. Neurophysiol.* **101**, 1813–1822 (2009).
- Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7**, 358–366 (2006).
- Roelfsema, P. R. Solving the binding problem: assemblies form when neurons enhance their firing rate—they don’t need to oscillate or synchronize. *Neuron* **111**, 1003–1019 (2023).

42. Self, M. W., van Kerkoerle, T., Super, H. & Roelfsema, P. R. Distinct roles of the cortical layers of area V1 in figure-ground segregation. *Curr. Biol.* **23**, 2121–2129 (2013).
43. Schiller, P. H. Effect of lesions in visual cortical area V4 on the recognition of transformed objects. *Nature* **376**, 342–344 (1995).
44. White, B. J., Kan, J. Y., Levy, R., Itti, L. & Munoz, D. P. Superior colliculus encodes visual saliency before the primary visual cortex. *Proc. Natl. Acad. Sci. USA* **114**, 9451–9456 (2017).
45. Lamme, V. A., Supér, H. & Spekreijse, H. Feedforward, horizontal, and feedback processing in the visual cortex. *Curr. Opin. Neurobiol.* **8**, 529–535 (1998).
46. Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H. & Bosch, H. Figure-ground segregation in a recurrent network architecture. *J. Cogn. Neurosci.* **14**, 525–537 (2002).
47. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 (1962).
48. Ferster, D., Chung, S. & Wheat, H. Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature* **380**, 249–252 (1996).
49. Gilbert, C. D. & Wiesel, T. N. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* **9**, 2432–2442 (1989).
50. Bosking, W. H., Zhang, Y., Schofield, B. & Fitzpatrick, D. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* **17**, 2112–2127 (1997).
51. Epshtein, B., Lifshitz, I. & Ullman, S. Image interpretation by a single bottom-up top-down cycle. *Proc. Natl. Acad. Sci. USA* **105**, 14298–14303 (2008).
52. Liu, Y. J. et al. Tracing inputs to inhibitory or excitatory neurons of mouse and cat visual cortex with a targeted rabies virus. *Curr. Biol.* **23**, 1746–1755 (2013).
53. Shen, S. et al. Distinct organization of two cortico-cortical feedback pathways. *Nat. Commun.* **13**, 6389 (2022).
54. Pfeffer, C. K., Xue, M., He, M., Huang, Z. J. & Scanziani, M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* **16**, 1068–1076 (2013).
55. Keller, A. J. et al. A disinhibitory circuit for contextual modulation in primary visual cortex. *Neuron* **108**, 1181–1193.e1188 (2020).
56. Hupe, J. M., James, A. C., Girard, P. & Bullier, J. Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *J. Neurophysiol.* **85**, 146–163 (2001).
57. Gilbert, C. D. & Li, W. Top-down influences on visual processing. *Nat. Rev. Neurosci.* **14**, 350–363 (2013).
58. Roelfsema, P. R. & de Lange, F. P. Early visual cortex as a multiscale cognitive blackboard. *Annu. Rev. Vis. Sci.* **2**, 131–151 (2016).
59. Maunsell, J. H. R. Neuronal mechanisms of visual attention. *Annu. Rev. Vis. Sci.* **1**, 373–391 (2015).
60. Bichot, N. P., Heard, M. T., DeGennaro, E. M. & Desimone, R. A source for feature-based attention in the prefrontal cortex. *Neuron* **88**, 832–844 (2016).
61. Moore, T. & Zirnsak, M. Neural mechanisms of selective visual attention. *Annu. Rev. Psychol.* **68**, 47–72 (2017).
62. Fiebelkorn, I. C. & Kastner, S. A rhythmic theory of attention. *Trends Cogn. Sci.* **23**, 87–101 (2019).
63. Martinez-Trujillo, J. Visual attention in the prefrontal cortex. *Annu. Rev. Vis. Sci.* **8**, 407–425 (2022).
64. Zhao, X. N. et al. Population coding for figure-ground texture segregation in macaque V1 and V4. *Prog. Neurobiol.* **240**, 102655 (2024).
65. Hembrock-Short, J. R., Mock, V. L. & Briggs, F. Attentional modulation of neuronal activity depends on neuronal feature selectivity. *Curr. Biol.* **27**, 1878–1887 (2017).
66. Chen, Y. et al. Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nat. Neurosci.* **11**, 974–982 (2008).
67. Martinez-Trujillo, J. C. & Treue, S. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr. Biol.* **14**, 744–751 (2004).
68. Vanduffel, W., Tootell, R. B. H. & Orban, G. A. Attention-dependent suppression of metabolic activity in the early stages of the macaque visual system. *Cereb. Cortex* **10**, 109–126 (2000).
69. Hopf, J.-M. et al. Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proc. Natl. Acad. Sci. USA* **103**, 1053–1058 (2006).
70. Li, W., Piech, V. & Gilbert, C. D. Learning to link visual contours. *Neuron* **57**, 442–451 (2008).
71. Sigman, M. et al. Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron* **46**, 823–835 (2005).
72. Lamme, V. A. The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15**, 1605–1615 (1995).
73. Gegenfurtner, K. R., Kiper, D. C. & Fenstemaker, S. B. Processing of color, form, and motion in macaque area V2. *Vis. Neurosci.* **13**, 161–172 (1996).
74. Mitzdorf, U. Current source-density method and application in cat cerebral cortex—investigation of evoked-potentials and EEG phenomena. *Physiol. Rev.* **65**, 37–100 (1985).
75. Schroeder, C. E., Mehta, A. D. & Givre, S. J. A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb. Cortex* **8**, 575–592 (1998).
76. Nandy, A. S., Nassi, J. J. & Reynolds, J. H. Laminar organization of attentional modulation in macaque visual area V4. *Neuron* **93**, 235–246 (2017).
77. Matsuda, K., Nagami, T., Sugase, Y., Takemura, A. & Kawano, K. A Widely Applicable Real-time Mono/Binocular Eye Tracking System using a High Frame-rate Digital Camera (Springer International Publishing, 2017).
78. Hautus, M. J. Corrections for extreme proportions and their biasing effects on estimated values of d' . *Behav. Res. Meth. Ins. C.* **27**, 46–51 (1995).
79. Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).
80. Gu, Y. et al. Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* **71**, 750–761 (2011).
81. Fisher, R. A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**, 179–188 (1936).

Acknowledgements

We thank Xibin Xu and Huaixue Chen for technical assistance. This work was supported by National Natural Science Foundation of China (31930049 to W.L., 31970929 to Y.Y.), STI2030-Major Projects (2022ZD0208200 to Y.Y., 2022ZD0204600 to W.L.), and Fundamental Research Funds for the Central Universities (2243300006 to Y.Y.).

Author contributions

Y.X., Y.Y., and W.L. designed research; Y.X. and Y.Y. performed research; Y.X. and Y.Y. analyzed data; Y.X., Y.Y., and W.L. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-62279-8>.

Correspondence and requests for materials should be addressed to Yin Yan or Wu Li.

Peer review information *Nature Communications* thanks Pieter Roelfsema and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025