# Opposing plasticity mechanisms in single neurons shape visual saliency assignment

Koen Seignette[1,2*], Leander de Kraker[1], Paolo Papale[3], Lucy S. Petro[4,5], Jorrit S. Montijn[6], Matthew W. Self[3,7*], Matthew E. Larkum[2], Pieter R. Roelfsema[3,8,9,10], Lars F. Muckli[4,5] & Christiaan N. Levelt[1,11,**]

1. Department of Molecular Visual Plasticity, Netherlands Institute for Neuroscience, Amsterdam, Netherlands
2. Department of Biology, Humboldt Universität zu Berlin, Berlin, Germany
3. Department of Vision & Cognition, Netherlands Institute for Neuroscience, Amsterdam, Netherlands
4. Centre for Cognitive Neuroimaging, School of Psychology and Neuroscience, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK
5. Imaging Centre for Excellence (ICE), College of Medical, Veterinary and Life Sciences, University of Glasgow, UK
6. Department of Cortical Structure & Function, Netherlands Institute for Neuroscience, Amsterdam, Netherlands
7. College of Medical, Veterinary and Life Sciences, University of Glasgow, UK
8. Laboratory of Visual Brain Therapy, Sorbonne Université, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Institut de la Vision, Paris, France
9. Department of Integrative Neurophysiology, Centre for Neurogenomics and Cognitive Research, VU University, Amsterdam, Netherlands
10. Department of Neurosurgery, Academic University Medical Center, University of Amsterdam, Amsterdam, Netherlands
11. Department of Molecular and Cellular Neurobiology, Center for Neurogenomics and Cognitive Research, VU University Amsterdam, Amsterdam, Netherlands

*Current affiliation

**Correspondence: (c.levelt@nin.knaw.nl)

**ABSTRACT**

To navigate complex environments, the visual system must prioritize unexpected inputs or omissions over predictable sensory backgrounds. This saliency assignment is a hallmark of visual processing and depends on comparing visual feedforward inputs with learned contextual surrounds, a process often conceptualized as a difference computation. However, cortical circuitry poses fundamental challenges: visual contextual feedback is predominantly amplifying, and feedback-driven inhibitory neurons necessary for subtractive computations are relatively sparse. How then can complex contextual feedback be selectively compared with feedforward input at the level of individual neurons? Here, we show that divergent plasticity rules within single pyramidal neurons provide a solution to this problem. Using two-photon calcium imaging, we recorded layer 2/3 pyramidal cell responses in mouse primary visual cortex before and after visual familiarization. We dissociated feedforward and contextual signals by comparing responses to stimuli with and without an occluded sector. Repeated feedforward responses selectively weakened through adaptation and surround suppression, whereas contextual responses strengthened and generalized across scenes. These strengthened contextual inputs amplified non-adapted responses to novel images and drove neuronal activity when occlusion of feedforward input prevented surround suppression. This produced a conserved structure of feedforward and contextual responses across mice, monkeys, and humans. Ultimately, these results reveal a cellular mechanism for saliency assignment that does not require feedback-driven inhibition. Pyramidal neurons autonomously learn the generalized statistical structure of familiar contexts, allowing them to selectively amplify improbable inputs or omissions.

## INTRODUCTION

Visual scenes contain far more information than can be processed rapidly. The visual system must therefore prioritize the most informative elements - features that deviate from their surroundings and learned contextual expectations[1–4]. Identifying these salient features requires comparing incoming feedforward signals with contextual information derived from past experience and the surrounding visual structure. How cortical circuits implement such comparisons remains a central challenge in neuroscience research[5–7]. Theoretical accounts often propose that deviations are detected through subtractive or difference-like computations between feedforward input and contextual feedback[8–13]. This mechanism may indeed account for attenuation of self-motion related responses[14]. However, the precise subtractive interactions required for computing unexpected inputs in complex visual scenes are not supported by the known functional anatomy of visual cortex, where contextual feedback is predominantly amplifying and inhibitory neurons are relatively sparse[15]. This raises a fundamental question: how does cortex compare feedforward visual input with complex contextual signals if both inputs are primarily excitatory?

Addressing this question requires experimentally dissociating feedforward sensory drive from contextual influences. In primary visual cortex (V1), neurons respond not only to stimuli within their classical receptive fields (cRFs), but also to contextual signals conveyed by recurrent and feedback connections[16], giving rise to extraclassical receptive field (ecRF) responses[17–20]. Partially occluded visual stimuli provide a powerful approach to disentangle these inputs as they eliminate feedforward input to a retinotopic region of V1 while preserving surrounding visual context[21]. Human fMRI studies[22–26] and electrophysiological recordings in non-human primates[27] have shown that occluded regions can evoke robust, image-specific responses in V1, driven by information from beyond the classical receptive field. These responses reflect contextual inputs rather than modulation of direct sensory input, making occlusion paradigms well suited for isolating contextual signals and studying their experience-dependent modulation.

Here, we used two-photon calcium imaging to record the activity of layer 2/3 pyramidal cells (PyCs) in mouse V1 during presentation of non-occluded and partially occluded natural images,

before and after visual familiarization. This approach allowed us to ask whether experience alters feedforward and contextual inputs differently at the level of individual neurons, and whether such changes can account for experience-dependent assignment of saliency to sensory inputs. Our results support a circuit mechanism for saliency assignment in which PyC connectivity is dynamically optimized to emphasize contrast between feedforward and contextual inputs without the need for feedback-driven inhibition.

## RESULTS

Studying the impact of visual experience on feedforward and contextual responses

To assess neural responses to non-occluded and occluded scenes, we recorded calcium signals from L2/3 PyCs in V1 of awake mice using wide-field (WF) and two-photon (2P) imaging while presenting six different natural images in both conditions (Fig. 1A–C). The luminance of the occluded area was equal to that of the gray screen during the intertrial interval, so that no change occurred in the cRF of the recorded neurons when the occluded image appeared (Fig. 1A). Mice were naive to visual input other than normal visual experience within the animal facility.

To test the effect of visual experience on these responses, we familiarized the mice in a simple stimulus detection task, allowing us to compare neural responses in naive and experienced animals. During behavioral training, we presented only four out of the six images to the mice, and only their non-occluded versions. To ensure engagement, mice received a small water reward if they licked within the time window 1–2 s after disappearance of any image. Mice learned to associate the stimuli with a reward, leading to anticipatory-licking and high hit rates (Fig. S1A-B). During passive imaging sessions before ("naive mice") and after ("expert mice") familiarization, mice were not performing the task but passively viewed the stimuli. To test the influence of engagement, we recorded responses in one extra 2P session during task performance ("during task"). Excluding two out of six images from the training paradigm allowed us to compare experience-dependent changes in neuronal responses to "familiar" (i.e. used for training) and "novel" (i.e. not used for training) images.

We performed calcium imaging in L2/3 PyCs by recording in TIT2L-GCaMP6f x G35-3-Cre mice, in which Cre is expressed in cortical excitatory neurons[28] (Fig. 1B). We targeted our recordings to the region of V1 that retinotopically represented the occluded quadrant of the natural scenes. To achieve this, we first used WF calcium imaging to perform a population receptive field (pRF) mapping[29] session immediately followed by the occlusion experiment (Figs. 1D-E and S1C).

We used the pRF maps to extract WF signals from V1 that represented the occluded quadrant of the scenes (taking a safety margin of 7 visual degrees to minimize bleed-through from neighboring regions) (Fig. S1C, E). The non-occluded and occluded scenes elicited robust

responses in all mice (Figs. 1D-E and S1C). This confirms that PyCs in mouse V1 representing the occluded region of a natural scene – and thus receiving no feedforward input from the dorsal lateral geniculate nucleus (dLGN) – are nevertheless activated by the occluded images, just like in humans[23,25] and monkeys[27].

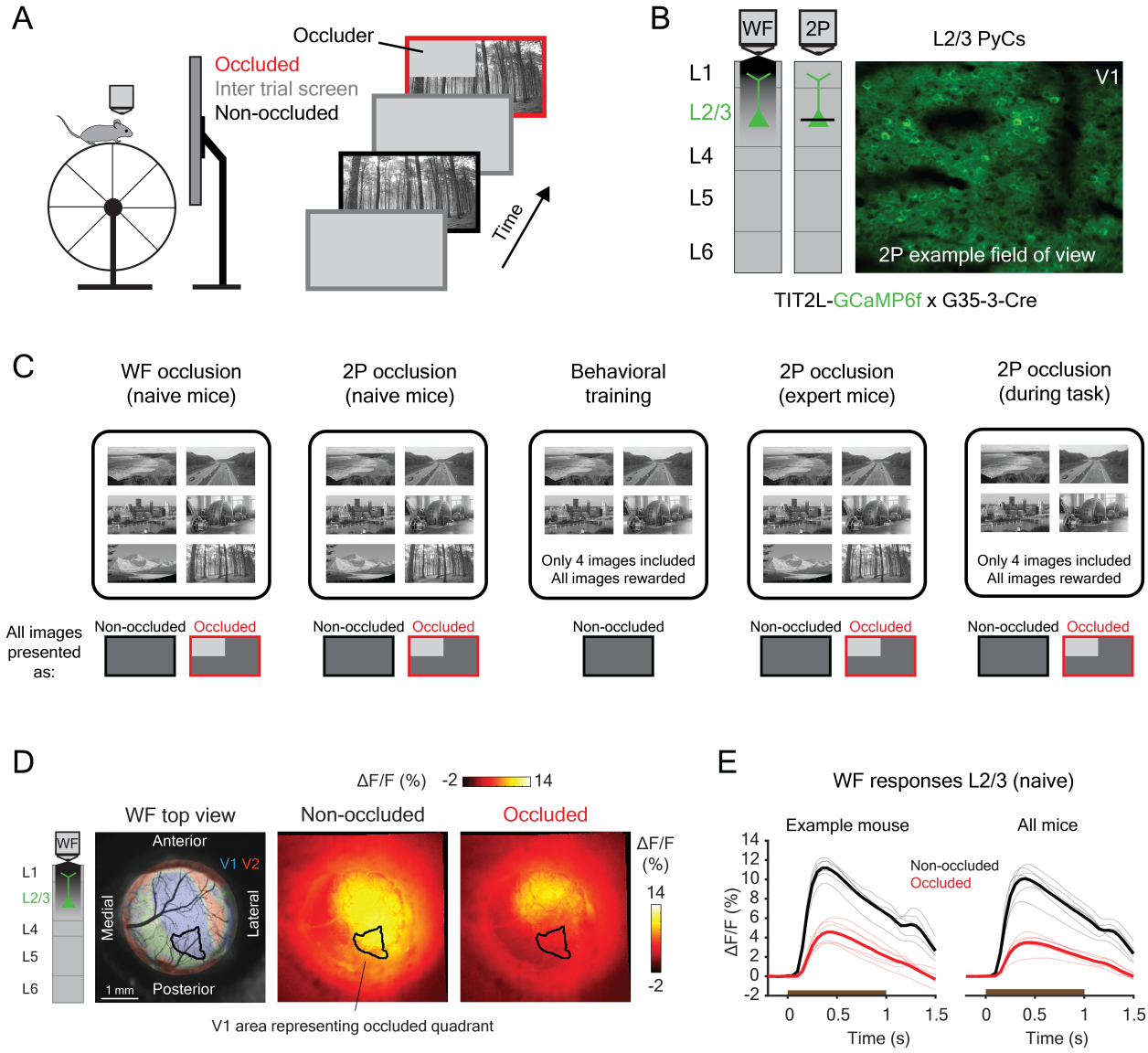


Figure 1. Experimental setup to study plasticity of visual and contextual responses across cell types in mouse V1

A) Schematic of experimental setup. Mice were head-fixed on a running wheel centered in front of a screen displaying non-occluded and occluded natural scenes.

B) Imaging approach showing wide-field (WF) and two-photon (2P) calcium imaging in L2/3 PyCs in V1 of TIT2L-GCaMP6f x G35-3 mice.

C) Schematic of calcium imaging and behavioral training steps. 'Naive' mice were subjected to a WF population RF mapping session followed by retinotopically targeted WF and 2P occlusion experiments. 2P sessions

were performed before (naive) and after (expert) a behavioral training phase to familiarize mice with the natural scenes. For behavioral training we used a subset (4/6) of the images, without occlusion. One additional 2P imaging session with familiar images was performed while the mice were engaged in the task (during task).

D) Cortical maps showing average ΔF/F responses to the non-occluded and occluded stimuli for an example mouse. From left to right: WF top view with V1 and V2 areas indicated in blue and red, response to non-occluded stimuli, response to occluded stimuli. Outlined in black is the area of V1 that represents the visual field corresponding to the occluded quadrant on the screen.

E) Average WF responses to non-occluded and occluded visual stimuli for an example mouse (left, thin lines represent individual images, thick line represents mean across images) and all mice (right, thin lines represent individual mice, thick line represents mean across mice).

## Experience and engagement drive plasticity of responses to non-occluded and occluded images in opposite directions

We used 2P calcium imaging to record the responses of individual L2/3 PyCs. We studied how these responses depended on visual experience by recording the activity of neurons before and after familiarization with the non-occluded versions of the natural scenes, and during task performance.

Using the WF pRF map, we targeted our 2P imaging to the region of V1 representing the occluder (Fig. 2A). We then mapped the cRF locations of all recorded neurons using sparse noise stimuli and applied strict criteria to only include neurons in all subsequent analyses with high quality cRFs falling on the occluded region and not on the non-occluded image regions (see Materials and Methods and Fig. 2B). This ensures that responses of these neurons to occluded images are caused by the visual context (ecRF input).

We first focused on responses of PyCs to familiar images that were repeatedly presented during the training phase (Fig. 2C-E). For statistical comparisons we used a linear mixed effects model (LMEM), which accounts for dependencies in the data and variance across mice. Visual experience had profound effects on PyC responses to non-occluded and occluded stimuli. In naive mice, many neurons responded to occluded stimuli (Fig. 2C), although responses to non-occluded stimuli were stronger. Notably, in expert mice, the magnitude of responses to occluded stimuli was strongly increased (Fig. 2D-E), an effect that was even more profound during task

performance. In contrast, responses to non-occluded stimuli showed the opposite pattern. They decreased in expert mice and mice performing the task.

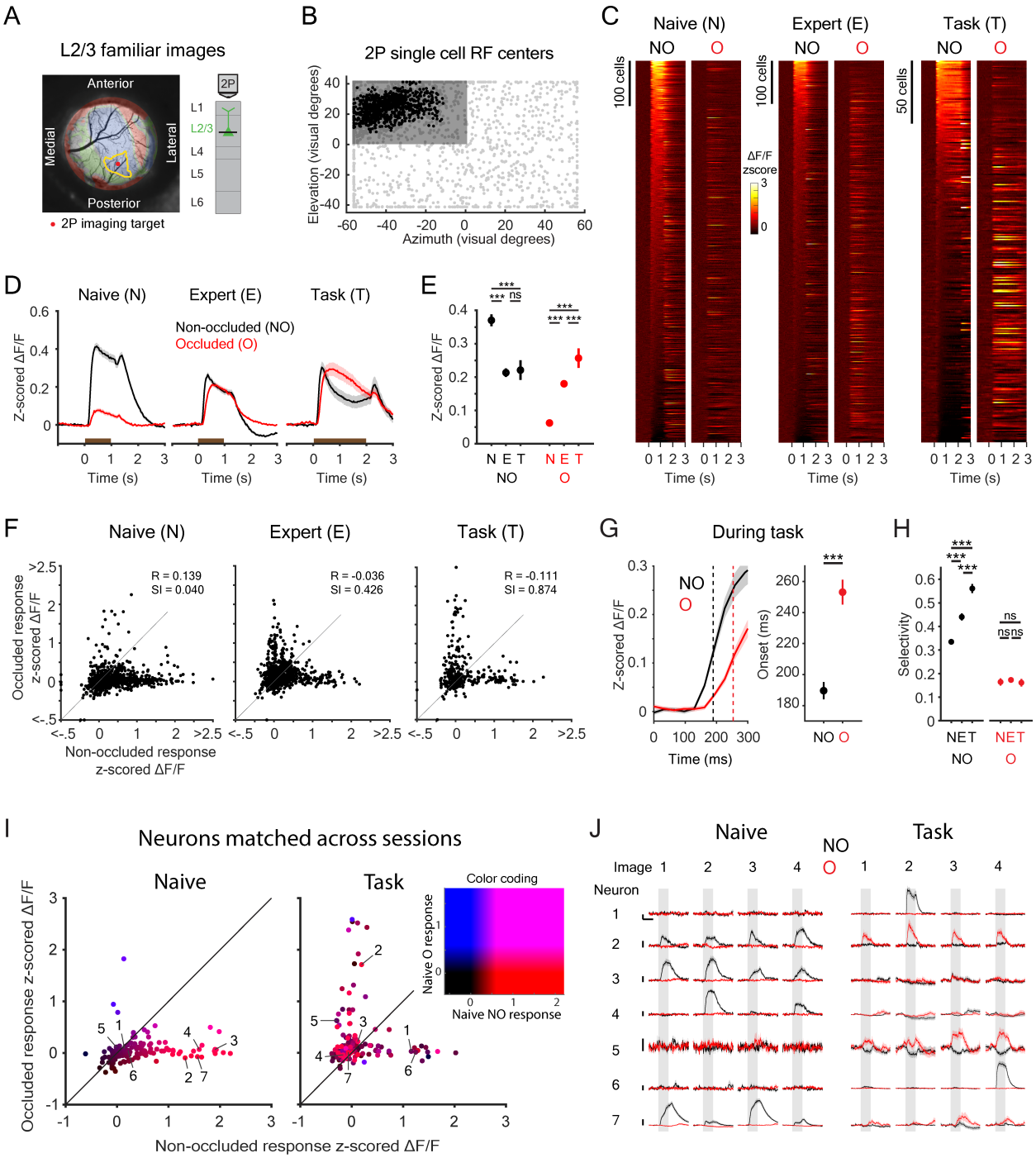

Figure 2. Experience shapes L2/3 responses to familiar feedforward and contextual inputs.

A) WF field-sign map showing V1 (blue) and higher visual areas (red) as obtained by population receptive field mapping for an example mouse. Green areas represent regions for which the sign could not be determined. The yellow outline indicates the area in V1 that represents the occluded quadrant on the screen. 2P imaging was targeted in the center of this region (red dot).
B) Visual field representation of the stimulus screen showing single cell RF centers. Gray dots are excluded RFs and black dots included RFs.
C) Heatmap of z-scored ΔF/F responses for individual neurons across conditions and stimuli. Neurons are sorted within each condition (Naive/Expert/Task) based on their average response (t = 0.2-1 s) to the non-occluded stimuli. N = 841 neurons (Naive), 884 neurons (Expert) and 300 neurons (Task) from 6 mice. NO: non-occluded, O: occluded.
D) Z-scored ΔF/F traces averaged across neurons and stimuli for all conditions. Traces represent mean ± SEM over neurons. Brown bars indicate stimulus epoch.
E) Mean response strength across stimuli and conditions (mean ± SEM, averaged between t = 0.2-1 s). LMEM with post hoc Tukey HSD for all comparisons, ***: p<0.001, ns: not significant.
F) Scatter plot showing average z-scored ΔF/F responses (t = 0.2-1 s) to the non-occluded (x-axis) and occluded (y-axis) stimulus across conditions. N = 841 neurons (Naive), 884 neurons (Expert) and 300 neurons (Task) from 6 mice. R, correlation coefficient; SI, separation index.
G) Left: zoom of mean response to NO and O images across neurons during task performance (mean ± SEM). Dotted lines represent mean onset times. Right: quantification of onset times (NO: n = 142 neurons, O: n = 174 neurons). LMEM for all comparisons, ***: p<0.001.
H) Stimulus selectivity (lifetime sparseness) for responses to NO and O stimuli. NO: N = 382 neurons (Naive), 223 neurons (Expert) and 79 neurons (Task) from 6 mice. O: N = 79 neurons (Naive), 182 neurons (Expert) and 79 neurons (Task) from 6 mice. LMEM with post hoc Tukey HSD for all comparisons, ***: p<0.001, ns: not significant.
I) Scatter plot showing average z-scored ΔF/F responses (t = 0.2-1 s) to the non-occluded (x-axis) and occluded (y-axis) stimulus of chronically matched neurons in naive mice and mice performing the task, color coded for response strength in naive mice. N = 133 neurons from 6 mice. Numbers of the indicated neurons correspond to the neurons in J.
J) Neuronal responses of individual, chronically matched example neurons in naive mice (left) and mice performing the task (right). NO, non-occluded; O, occluded. Scale bars, 1 s, 1 z-scored ΔF/F. Traces represent mean ± SEM over trials. The shaded bars indicate stimulus epochs.

As locomotion is known to modulate visual responses in mouse V1[18,30], we assessed whether changes in locomotion influenced the observed plasticity of responses to non-occluded and occluded stimuli but found that this was not the case (Fig. S2A-B). We also ruled out the possibility that mice made eye movements in response to stimulus presentations, which might otherwise have positioned non-occluded image regions in the cRFs (Fig. S2C-D).

We conclude that visual experience decreased responses of L2/3 PyCs to non-occluded images, while it increased responses to the visual context provided by occluded images. Task engagement

further enhanced responses to occluded images, suggesting that it strengthens responses to the visual contextual inputs.

Separation of L2/3 PyCs responding to feedforward and contextual information

We next analyzed how responses to non-occluded and occluded images were distributed across neurons. Interestingly, many PyCs appeared to respond to either the non-occluded stimuli or the occluded stimuli but not to both. This separation strengthened after familiarization with the images ("expert") and was most pronounced in the condition under which the mice had been familiarized ("task") (Fig. 2C, F, for responses to pairs of individual occluded and/or non-occluded images see Figs. S3 and S4). To quantify the extent of this unexpected separation of PyCs into two non-overlapping populations, we computed a separation index (SI) ranging from -1 to 1 (see Materials and Methods). The SI is positive when the degree of separation is higher than expected under the assumption that activity elicited by non-occluded and occluded stimuli is uncorrelated across PyCs. The SI is negative when both categories of stimuli activate the same neurons. The SI increased with experience and task engagement (Fig. 2F, SI: 0.040 for naive, 0.426 for expert and 0.874 for task). In parallel, the correlation coefficient (R) between population responses to non-occluded and occluded stimuli (averaged across images per neuron) decreased (R: 0.139 for naive, -0.036 for expert, -0.111 for task), affirming the separation in responses to the two types of stimuli.

The responses of PyCs to occluded images were reminiscent of inverse receptive field responses, which have been shown to depend on feedback inputs from higher order visual areas[31,32]. As responses to feedback inputs are characterized by a delay in response onset[27,31,32], the onset latency of calcium responses to occluded images is expected to be longer than to non-occluded images if they are driven by feedback inputs. We determined latencies by fitting a curve to the activity time course and taking the time point at which the curve reached 33% of its maximum[32]. We only included neurons with high quality fits (see Materials and Methods). Although calcium activity rises slower than spiking activity, the responses elicited by occluded stimuli were indeed delayed compared to responses elicited by non-occluded images (Fig. 2G, 189 ± 5.6 ms for non-

occluded and 253 ± 8.0 ms for occluded, $P < 0.001$, LMEM). This delayed response is in line with the interpretation that they were predominantly driven by feedback inputs, but does not rule out a contribution of lateral connections within V1.

We tested whether the responses to occluded stimuli could be explained by responsiveness to the nearby edge of the occluded region, in which case the responsiveness would depend on proximity to this edge. However, the correlation between response strength and cRF distance from the edge of the occluded region was weak and statistically non-significant (Fig. S5, R = -0.005, $P = 0.876$).

After familiarization with non-occluded natural images, PyCs became more selective for individual images (Fig. 2H; lifetime sparseness). This increase in selectivity was associated with preferential suppression of neurons that responded to multiple images, as stimulus selectivity in naive mice was negatively correlated with the change in response amplitude after familiarization (R = −0.299, P = 0.001; Fig. S6). Selectivity for occluded images was much lower and did not increase with experience (Fig. 2H).

Taken together, our results indicate that responses to feedforward and contextual inputs predominantly occur in separate populations of L2/3 PyCs. Moreover, familiarization weakens responses to feedforward inputs, mostly affecting PyCs responding to multiple images. In contrast, responses to contextual inputs increase upon familiarization, in line with previous work[33].

<u>Plasticity of feedforward and contextual responses occurs in opposite directions in individual neurons</u>

To understand how these changes at the population level were caused by changes in response properties of individual neurons, we matched neurons recorded at the different stages of the experiment and assessed how their responses to non-occluded and occluded images changed. This revealed that many PyCs responding to non-occluded images in naive mice mostly became unresponsive in expert mice (Fig. S7), especially when engaged in the task (Fig. 2I, red cells

unresponsive in the “task” panel; Fig. 2J, cells 3, 4 & 7). In contrast, a subset of PyCs that were initially unresponsive became responsive to occluded images after familiarization (Fig. 2I, black/purple cells responsive to occluded images in the “task” panel; Fig. 2J, cell 5).

Intriguingly, some PyCs that were initially responsive to non-occluded images had become responsive to the occluded images (Fig. 2I, red cells responsive to occluded images in the “task” panel; Fig. 2J, cell 2). Although individual PyCs initially responding to non-occluded images could become responsive to occluded ones, they never responded strongly to both image types. Finally, a subset of neurons that was not or weakly responsive to the familiar images in naive mice enhanced their responses to the same images in expert mice (Fig. S7), and in particular during task engagement (Fig. 2I, black/purple cells responsive to non-occluded images in the “task” panel; Fig. 2J, cells 1 & 6). These cells were highly selective for the individual familiarized images (lifetime sparseness of 0.59).

We conclude that also at the level of individual L2/3 PyCs, repeated visual stimulation drives feedforward and contextual plasticity in opposite directions. This reorganization can shift cells from unresponsive or feedforward-driven toward strongly context-driven. In a small subset of PyCs, however, selective strengthening of feedforward responses occurs. Strikingly, strong feedforward and contextual responses are never expressed within the same cell.

## Separation of L2/3 PyCs responding to occluded and non-occluded images is caused by differences in extraclassical receptive field inputs and surround suppression

We next sought to understand the functional basis of the separation of PyCs responding to feedforward and contextual inputs. We reasoned that PyCs responding to occluded images may receive strong ecRF inputs, while being suppressed during presentation of non-occluded images. Such suppression could arise from local inhibitory mechanisms engaged when nearby PyCs with overlapping or adjacent cRF are co-activated, a phenomenon known as surround suppression (Fig. S8). Importantly, this form of suppression would be minimal for occluded images, as well as for the sparse noise stimuli used to map cRFs.

To test this hypothesis, we repeated the experiment with additional stimuli, including gratings in the form of discs and annuli of varying sizes to assess surround suppression and responses to stimuli beyond their cRFs, respectively, both before and after familiarizing mice with the non-occluded natural images. To accelerate the familiarization procedure, we slightly changed the paradigm. The mice now only passively viewed 4 of the 6 non-occluded images during 6 sessions across 3 days without rewards or lick training (Fig. 3A). After familiarization, we again observed that distinct L2/3 PyC populations responded to occluded and non-occluded images (Fig. 3B). Familiarization weakened responses to non-occluded images, while responses to occluded images were increased, albeit slightly less strongly than after reward-assisted training (Fig. 3C).

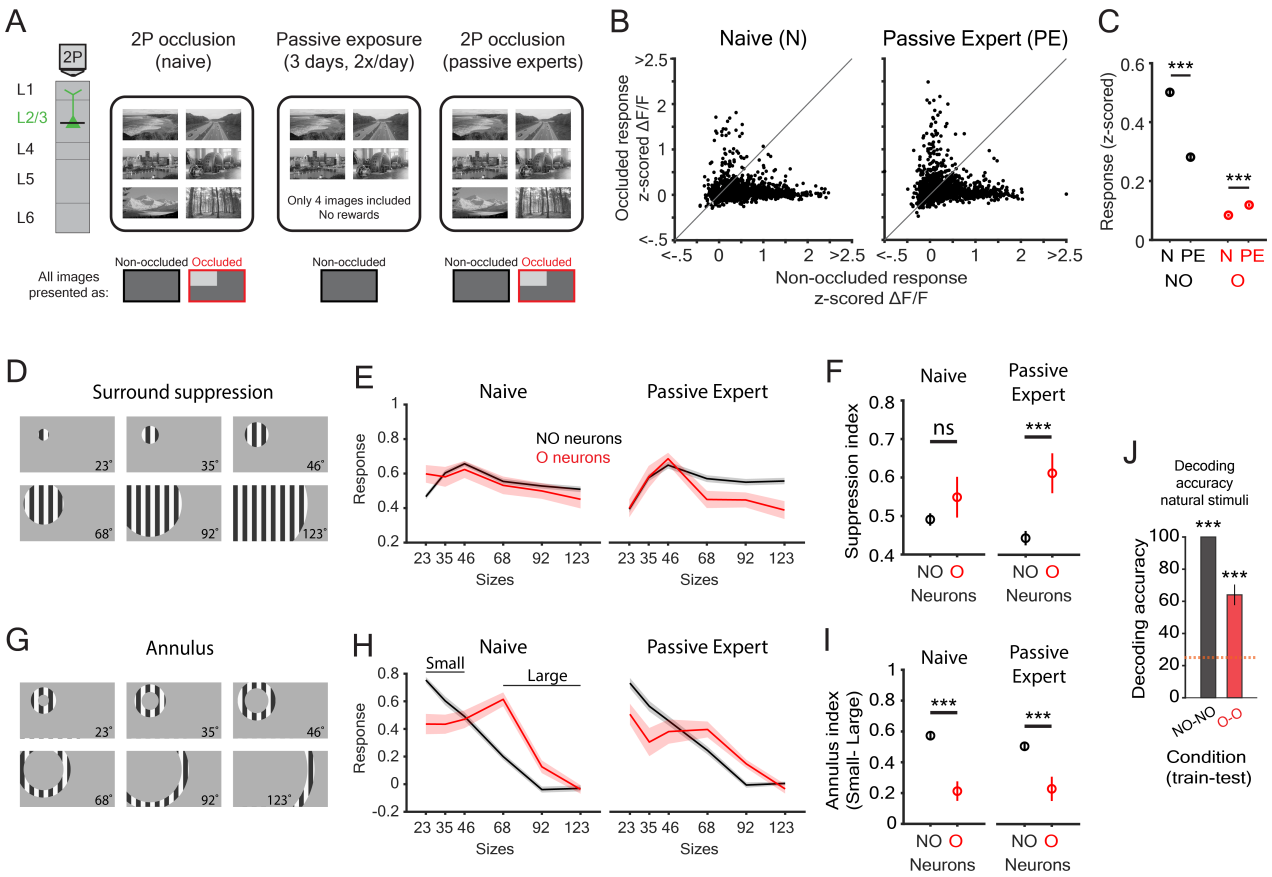


Figure 3. Response properties of L2/3 neurons responding to NO and O images.

A) Schematic of experimental setup and passive exposure training regime. Mice passively viewed 4 of the 6 non-occluded images during 6 sessions in 3 days without lick training or rewards.

B) Scatter plot showing average z-scored ΔF/F responses (t = 0.2-1 s) to the non-occluded (x-axis) and occluded (y-axis) stimulus across conditions. N = 1820 neurons (Naive) and 2106 neurons (Passive Expert) from 7 mice.
C) Mean response strength across stimuli and conditions (mean ± SEM, averaged between t = 0.2-1 s). LMEM for all comparisons, ***: p<0.001.
D) Surround suppression stimuli with their respective sizes in visual degrees as used in this experiment.
E) Average surround suppression for neurons responding to NO or O stimuli (response > 0.5) in Naive and Expert mice. Note the stronger surround suppression for neurons responding to O stimuli. Traces represent mean ± SEM over neurons.
F) Mean suppression index for NO vs O responders in Naive and Passive Expert mice (mean ± SEM, calculated as ($R_{max}$ - $R_{large}$) / $R_{max}$). NO neurons Naive: N = 691, O neurons Naive: N = 61. NO neurons Passive Expert: N = 392, O neurons Passive Expert: N = 100, from 7 mice.
G) Annulus stimuli with their respective sizes in visual degrees as used in this experiment. The inner circle of the annulus matched the diameter of the surround suppression stimuli as shown in D).
H) Annulus response curves for neurons responding to NO or O stimuli (response > 0.5) in Naive and Passive Expert mice. Note the stronger annulus responses at larger sizes for neurons responding to O stimuli. Traces represent mean ± SEM over neurons.
I) Mean annulus index for NO vs O responders in Naive and Expert mice (mean ± SEM, calculated as $R_{small}$ – $R_{large}$). NO neurons Naive: N = 691, O neurons Naive: N = 61. NO neurons Passive Expert: N = 392, O neurons Passive Expert: N = 100, from 7 mice.
J) Accuracy of population decoding of stimulus identity in Passive Expert mice. An LDA decoder was trained on separate sets of trials for cross-validation of non-occluded responses (NO-NO, train-test) and occluded responses (O-O). Bars show mean ± SEM across repetitions of decoding runs. Chance performance (orange dashed line) was 25%. N = 2106 neurons from 7 mice. Permutation test, ***: p<0.001.

In line with the explanation described above, after familiarization neurons responsive to occluded images showed significantly stronger surround suppression than those responsive to non-occluded images (Fig. 3D-F). Furthermore, neurons responsive to occluded images preferred larger annuli than those responsive to non-occluded images, and unlike the latter, they also responded to annuli positioned entirely outside the occluded quadrant (92°; Fig. 3G–I). This suggests that although occluded image-responders also received ecRF input from within the upper left quadrant, inputs from outside this region were sufficient to drive their responses to occluded images.

To assess whether responses to occluded images carried stimulus-specific information, we performed a population classification analysis using all neurons with cRFs within the occluded region, averaging responses for each neuron and each trial (between 0.2-1 s after stimulus onset). We used 50% of the trials to train a linear classifier (linear discriminant analysis; LDA) and the 50% remaining trials for testing. Training and testing the classifier on either non-occluded (Train-

Test, i.e. NO-NO) or occluded (O-O) responses allowed us to examine whether these signals contain stimulus-specific information (Fig. 3J). We found that the identity of non-occluded images could be decoded with 100% accuracy. Decoding accuracy of occluded images was lower (62%), but well above chance level (25%), confirming a limited degree of specificity of L2/3 PyCs responses to occluded images. Together, these results confirm that responses to occluded images arise from a combination of weakly selective contextual inputs from beyond the occluder in combination with release from surround suppression, possibly signaling a salient absence of a probable feedforward input.

Responses to novel, non-familiarized images

We next asked whether the salient presence of an improbable feedforward input would also evoke enhanced responses. We therefore examined how L2/3 PyCs responded to the two images the mice had not been familiarized with (Fig. 4, novel stimuli, based on the dataset obtained from our reward-assisted paradigm). Unlike the weakening of responses to familiar images, the activity elicited by these "novel" images had increased after training (Fig. 4A-D). Notably, this response increase was also present when occluded novel images were presented (Fig. 4A-D).

Given that responses to non-occluded images were more selective than responses to occluded images (Figs. 2H and 3J), we hypothesized that plasticity of these responses would follow the same pattern: repeated feedforward inputs would undergo selective weakening, whereas repeated ecRF inputs would be strengthened in a broader, more generalized way.

In this scenario, since novel images were not shown during training, responses selective for non-occluded novel images should not reduce. Moreover, if ecRF inputs strengthen in a manner that generalizes across stimuli, the contextual inputs recruited by novel images should overlap with those recruited by familiar images. This overlap would account for the enhanced responses to occluded novel images. Finally, because feedforward responses selective for non-occluded novel images remain intact, they could be further amplified by the strengthened contextual inputs, producing the increased responses we observe for non-occluded novel images.

We evaluated three predictions derived from these hypotheses. First, we assessed the overlap in contextual inputs for familiar and novel images, by correlating the responses of PyCs to occluded versions of the familiar and novel images. PyCs that responded strongly to occluded novel images in expert mice indeed also responded to occluded familiar images (Fig. 4E, R=0.789, P<0.001). Moreover, neurons that increased their responses to the occluded novel images also increased their responses to occluded familiar images (Fig. 4F, R=0.791, P<0.001). This confirms that strengthened ecRF responses generalize across natural images.

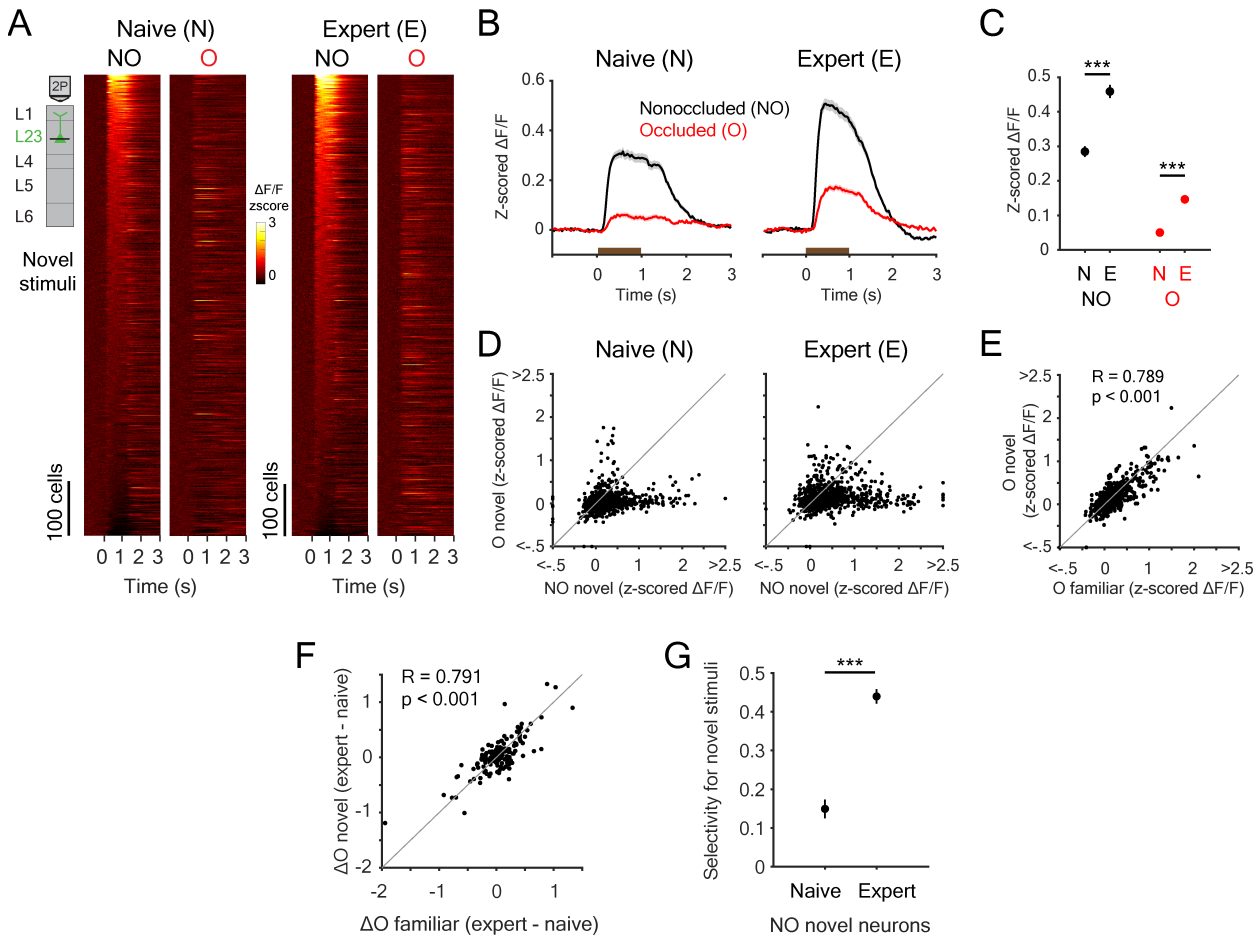


Figure 4. Novelty responses affect the most selective neurons and involve enhanced contextual inputs

A) Heatmap of z-scored ΔF/F responses for individual L2/3 neurons across conditions and stimuli for novel images. Neurons are sorted within each condition (N/E) based on their average response (t = 0.2-1 s) to the non-occluded stimuli. N = 841 neurons (Naive), 884 neurons (Expert) from 6 mice.

B) Z-scored ΔF/F traces averaged across neurons and stimuli for novel images. Brown bars indicate the visual stimulus epoch.
C) Mean response strength across stimuli and conditions (mean ± SEM, averaged between t = 0.2-1 s). LMEM for all comparisons, ***: $p<0.001$.
D) Scatter plot showing average z-scored ΔF/F responses to non-occluded novel stimuli (x-axis) and occluded novel stimuli (y-axis) in naive mice (left) and expert mice (right).
E) Scatter plot showing average z-scored ΔF/F responses to occluded familiar stimuli (x-axis) and occluded novel stimuli (y-axis) in expert mice.
F) Scatter plot showing increase in average z-scored ΔF/F responses to occluded familiar stimuli (x-axis) and occluded novel stimuli (y-axis) of chronically matched neurons (N = 166 neurons).
G) Selectivity for novel stimuli of responses to these stimuli in naive and expert mice. Included neurons responded to the novel NO image. Naive: N = 172 neurons, Expert: N = 285 neurons. Mean ± SEM, LMEM, $p<0.001$.

Second, we tested whether familiarization left responses to non-occluded novel images selectively intact. This is expected to increase selectivity for novel versus familiar images after familiarization. Indeed, we observed a significant experience-dependent increase in selectivity for novel over familiar images (Fig. 4G), consistent with a selective weakening of feedforward inputs.

Finally, if strengthened ecRF inputs amplify responses to novel images, PyCs responding to novel images should be strongly influenced by ecRF inputs. We therefore quantified the relative contribution of visual input from inside and outside the receptive field of the recorded neurons. We presented 3,600 natural images to the mice previously familiarized with the 4 non-occluded images through passive exposure, and computed the reverse correlations between the structure of the images and the neural responses. To characterize image structure, we used local Shannon entropy, which captures the complexity of pixel intensities within an image region. This allowed us to weight local image information by the strength of the neuronal responses it elicited. We then generated entropy-based activation maps for each neuron and averaged these maps for cells responsive to occluded and non-occluded, novel and familiar images (Fig. S9A). By comparing the ratio of activity elicited by inputs outside versus within the compound cRF region (RFs of all neurons together) across neuronal categories, we found that PyCs responding to novel non-occluded images received stronger ecRF inputs than those responding to familiar images (Fig. S9B). Together, these findings support the idea that enhanced novelty responses in L2/3 PyCs arise from non-adapted selective feedforward responses amplified by strengthened ecRF inputs that generalize across different natural scenes.

Non-occluded and occluded scenes elicited stimulus-specific neural signatures across species

Consistent with previous findings in humans and non-human primates, stimulus-specific activity in mouse V1 allowed reliable decoding of natural image identity from responses to both non-occluded and occluded stimuli[23,25,27]. We trained linear classifiers on population responses of PyCs and tested decoding performance for non-occluded images (Train-Test, NO–NO), occluded images (O–O), and across conditions (NO–O and O–NO; Fig. 5A and Fig. S10).

Decoding performance across conditions in mice was strikingly similar to that in humans and non-human primates (Fig. 5B-D), despite the differences in species and recording techniques. Across all species and recording modalities, decoding accuracy was highest within stimulus conditions and lower for cross-condition decoding. We observed high (>80%) and significant decoding of responses to non-occluded stimuli (NO-NO). O-O decoding accuracy was lower but significant, just like in humans and non-human primates (Fig. 5B). Finally, across all species and conditions, decoding accuracy was lower for cross-condition decoding, as expected given that responses to feedforward and contextual inputs predominantly arose from different neuronal populations. Interestingly, classifiers trained on occluded responses and tested on non-occluded responses (O-NO) were more accurate than the other way around. This asymmetry may reflect that classifiers trained on non-occluded images rely more strongly on feedforward responses, which are absent for occluded images, whereas classifiers trained on occluded images exploit contextual responses that also weakly contribute to non-occluded responses. We conclude that responses of PyCs to non-occluded and occluded images encode image-specific information and are highly similar across species.

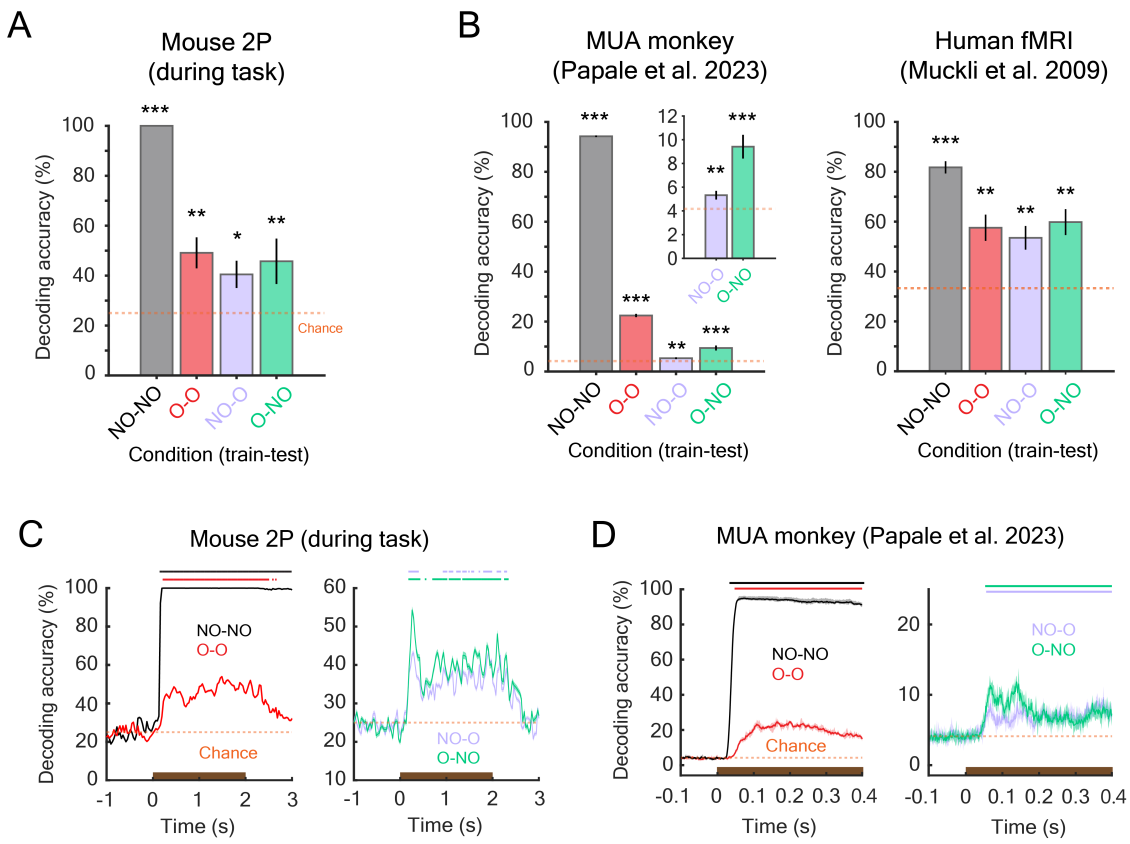


Figure 5. Responses to occluded scenes contain stimulus-specific information in mouse, monkey and human visual cortex.

A) Accuracy of population decoding of stimulus identity for mouse PyCs (left, n = 300 PyCs) during the task. An LDA decoder was trained and tested on separate sets of trials for cross-validation (NO-NO, O-O) or cross-decoding (NO-O, O-NO). Bars show mean ± SEM across repetitions of decoding runs. Chance performance (orange dashed line) was 25%. Permutation test, ***: p<0.001, **: p<0.01, *: p<0.05, ns: not significant. NO: non-occluded, O: occluded.

B) Left: decoding accuracy for multi-unit activity (MUA) monkey data. Chance performance was 4.16%. Permutation test, ***: p<0.001, **: p<0.01. Bars show mean ± SEM across repetitions of decoding runs. Inset shows a zoom for NO-O and O-NO conditions. Right: decoding accuracy for human fMRI data. Chance performance was 33.3%. Bars show mean ± SEM across individual subjects (n = 6). T-test against chance, ***: p<0.001, **: p<0.01.

C) Population decoding accuracy over time for mouse PyCs. Colored dashes at the top depict significant time points (p<0.05, permutation test). Brown bars indicate the time window with the visual stimulus. Orange dashed line represents chance performance.

D) As in C, but for monkey MUA data.

## DISCUSSION

Our results show that familiarization to natural images induces a striking rebalancing of feedforward and contextual influences on L2/3 PyCs in V1. Repeated feedforward responses are selectively weakened, whereas contextual responses strengthen and generalize across images. Acting together within individual neurons, these opposing plasticity mechanisms enable PyCs to extract salient signals and omissions from predictable visual backgrounds, without requiring contextual feedback-driven inhibition.

Several mechanisms may contribute to the downscaling of feedforward-driven neural activity after familiarization. These include synaptic depression[34,35] and reduced intrinsic excitability that both occur in relatively short time spans[36,37], and synaptic scaling via decreased postsynaptic AMPA receptor conductance over longer periods[38,39]. Additionally, repeated activation can strengthen inhibition, selectively suppressing highly active excitatory neurons[40,41] and contributing to surround suppression. Acting across different timescales and levels of organization, these mechanisms may synergistically reduce responses to familiar stimuli, while contextual influence is increasing.

Repeated presentation of natural visual scenes particularly reduced feedforward responses of less selective PyCs that responded to multiple images during familiarization. This disproportionate reduction of broadly-tuned neuron responses increased the relative contribution of more selective neurons, leading to a sharpening of visual responses, in accordance with previous work[42–44]. This sharpening through the plasticity of feedforward connections is not expected in classical formulations of hierarchical predictive coding that emphasize feedback-driven inhibition 'explaining away' expected inputs[42,45–49].

A subset of PyCs responded to occluded natural scenes despite the absence of cRF input, reminiscent of inverse receptive field responses driven by feedback inputs in the absence of feedforward drive[31]. This population of occluded image responders increased following familiarization with non-occluded scenes. Most of these neurons did not respond to the non-occluded familiar images, suggesting that strengthening of ecRF inputs does not require somatic spiking activity. This observation aligns with evidence that potentiation of feedback inputs onto

L2/3 PyC dendrites can occur independently of somatic firing[50] but may require dendritic NMDA spikes[51]. Disinhibition may have also supported the strengthening of ecRF inputs[32,52]. However, since most PyCs that responded to occluded images in expert mice did not at all respond to them before familiarization, disinhibition alone is unlikely to explain the observed changes. Possibly, disinhibition may have facilitated plasticity of ecRF inputs onto L2/3 pyramidal cells[53,54].

While PyC responses to different occluded images generalized across scenes, they nevertheless carried image-specific information, as stimulus identity could be decoded from these responses. Moreover, because PyCs responding to occluded images also received input from more proximal visual locations, as revealed by annulus stimuli, they appear to integrate spatially distributed contextual signals across the surround. Synaptic strengthening in apical dendrites depends on the co-activation of nearby synapses[50], suggesting that this form of plasticity could support the learning of structured contextual regularities rather than merely reinforcing frequently occurring features. Such contextual learning at the level of individual neurons may therefore enable selective amplification of inputs that violate these learned regularities. Future experiments will be needed to characterize the statistical structure of these contextual signals and to determine how it depends on stimulus proximity and spatial arrangement.

Occlusion-responsive neurons exhibited more surround suppression and received more distal ecRF inputs than those responding to non-occluded images. This combination of response properties may explain responses to occluded image responses, but also inverse receptive field responses[31,32]. Weakly selective inhibition from local interneurons, driven when neighboring neurons are activated by the non-occluded stimulus, suppresses occluded image responders[55]. When feedforward input near the cRF is absent, inhibition is reduced, allowing input from the surrounding visual context to drive the PyC's response (Fig. S8). These responses may highlight regions where visible input is missing from a scene, contributing to the perception of such areas as figures[32]. Whether such responses can also support contextual completion[56] when parts of familiar scenes are blurred or ambiguous is currently unclear.

Strengthening of ecRF inputs may also explain enhanced responses elicited by novel images presented after familiarization. Because these contextual inputs generalized across familiarized

and novel images, PyCs selective for these novel images likely receive stronger visual contextual inputs after familiarization. However, unlike PyCs responsive to familiar images, they did not undergo adaptation during familiarization. Consequently, the coincidence of non-adapted feedforward and strengthened contextual input may enhance feedforward drive in a context-dependent way[57]. This could involve apical amplification, consistent with models in which apical dendrites integrate contextual input and modulate somatic gain[58–61]. Supporting this idea, we observed that PyCs responding to novel images were more strongly influenced by inputs outside their cRFs than those responding to familiar images.

While most PyCs reduced their responses to familiar images, a small subset showed increased responsiveness and strong selectivity for familiarized stimuli. These neurons may not compute how feedforward inputs relate to context, but instead encode learned, behaviorally relevant signals. If so, they could represent a distinct subtype of L2/3 PyCs shaped by Hebbian plasticity, potentially enabled by feedback aligned with their cRF inputs. Consistent with this idea, it was found that L2/3 feedback aligns with cRF responses, whereas feedback from L5 PyCs in the lateromedial area originates from more distal ecRF regions and is not anatomically aligned with postsynaptic orientation preference[62].

Recent work showed that the emergence of such mismatched feedback during development can be modeled by Hebbian plasticity of feedforward inputs combined with anti-Hebbian plasticity of feedback inputs[63]. While this developmental plasticity also wires PyCs to enhance responses to deviations between feedforward and feedback signals, its functional outcome differs from what we observe in adult V1: the model predicts that experienced feedforward responses are strengthened, whereas matching feedback connections become disconnected.

Strikingly, decoding of non-occluded and occluded images, as well as cross-decoding between them, was similarly efficient across humans, monkeys, and mice, despite being measured with fMRI, electrophysiology, and 2P calcium imaging, respectively. This convergence across species and recording modalities suggests that common neural mechanisms underlie contextual processing, consistent with the circuit principles described above. The above-chance decoding of occluded natural images in naïve mice suggests that contextual inputs are already structured

before familiarization with these images, potentially reflecting connectivity shaped during development or by natural vision during adulthood.

The simplest circuit-level explanation consistent with our findings, is one in which two plasticity mechanisms with opposite directionality in individual PyCs dynamically optimize cortical responses to salient visual stimuli (Fig. 6). One mechanism selectively weakens responses to frequently active feedforward inputs[34,38,39,42,45,64–67], while the other strengthens regularly active contextual inputs[33,50,51,53] that generalize across scenes. This strengthening has two major consequences: 1) it amplifies responses to novel, non-adapted feedforward inputs that violate the statistical structure of familiarized contexts and 2) when local feedforward drive recruiting inhibition is absent, it enables contextual signals to drive responses that may signal the absence of expected inputs. By maximizing the contrast between feedforward and contextual influences within individual L2/3 PyCs, this circuitry provides a cellular mechanism for detecting and enhancing salient information..

This context-contrasting circuitry fits well with Helmholtz's classic idea of unconscious inference, in which perception reflects experience-based interpretation rather than passive sensory reception[68]. By amplifying responses to deviations from learned contextual structures, this circuitry may support rapid detection of behaviorally relevant stimuli within familiar scenes. When feedforward input is missing, strengthened context-driven responses may highlight such occluded regions, contributing to their perception as coherent shapes. These responses may also support contextual completion when inputs are degraded or ambiguous. Importantly, this concept may also help shed light on psychiatric symptoms such as psychosis: if the processes underlying context-contrasting plasticity are disrupted, the brain may assign saliency to inappropriate inputs, potentially contributing to hallucinations, delusions, or disorganized thinking[69]. By identifying a potential cellular mechanism for disrupted saliency processing within individual neurons, context-contrasting circuitry may offer new mechanistic approaches for understanding these conditions.

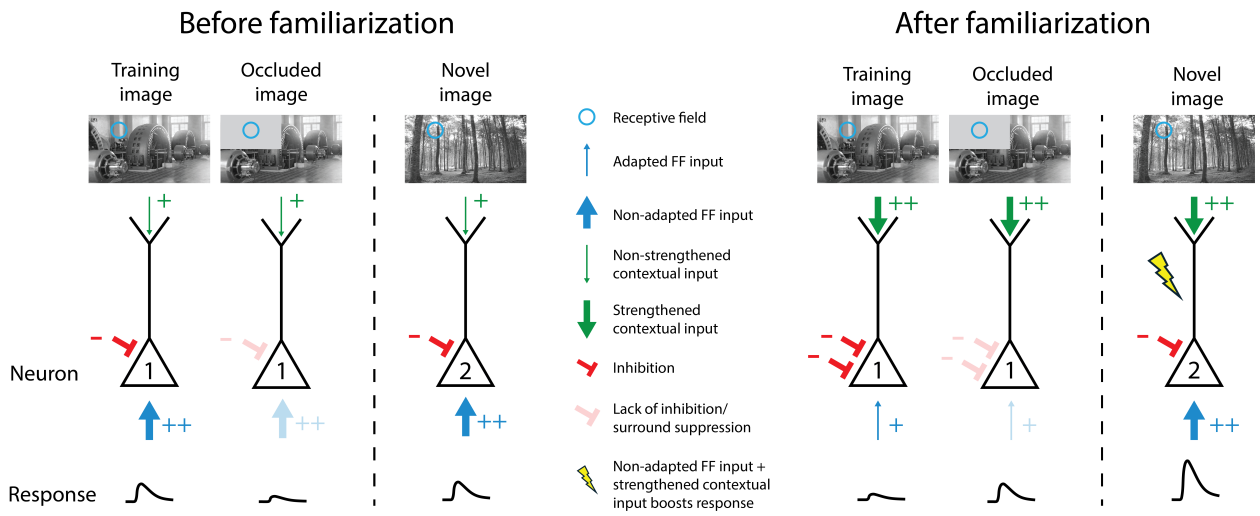


Figure 6. Divergent plasticity of feedforward and contextual inputs generates enhanced responses to salient stimuli.

Before familiarization (left), PyCs integrate feedforward (FF) visual input within their classical receptive fields and contextual input from surrounding regions. Non-occluded training- and novel images (neurons 1&2) evoke responses driven through FF inputs and modulated by context. Occluded images elicit little activity due to weak contextual drive and absence of FF input (neuron 1).

Upon familiarization (right), FF input from the training image selectively adapts and recruits stronger surround suppression (neuron 1), while contextual inputs strengthen in a generalizing way (neurons 1&2). As a result, responses to the non-occluded training image are reduced. In contrast, the occluded training image engages the strengthened contextual input in the absence of FF-driven inhibition, driving neural activity (neuron 1). The novel image also recruits this strengthened generalized context that enhances responses in PyCs that receive selective, non-adapted FF inputs from this image (neuron 2).

Together, these opposing plasticity processes selectively amplify improbable inputs or omissions by maximizing the contrast between feedforward and contextual signals in individual neurons.

## MATERIALS AND METHODS

### Mice

All experiments were approved by the institutional animal care and use committee of the Royal Netherlands Academy of Arts and Sciences. We used 4 female and 2 male G35-3Dtl[28] crossed to AI93 TITL-GCaMP6f mice (Jackson Laboratories, www.jaxmice.jax.org, strain 024107). Mice were group housed under a 12 h reversed day/night cycle and provided with ad libitum access to food and water. Experiments were performed in the dark phase. Mice had access to a running wheel in their home cage throughout the experiment.

### Window surgery and viral injections

For the cranial window surgery, mice were anesthetized with isoflurane (4% induction, 1.6% maintenance in oxygen). During the surgeries, temperature was maintained at 37$^{o}$C with euthermic pads and eyes were protected from light and drying using Cavasan eye ointment. Mice were implanted with a double (3+4 mm diameter) glass window and a custom metal head-ring on the center of V1 to allow head-fixation. The glass window and head-ring were fixed to the skull using dental cement. At least one week after the window surgery, mice were handled and trained to be head immobilized in our setup until they sat comfortably and started running regularly on their own.

### Two-photon imaging

Two-photon experiments were performed on a two-photon microscope (Neurolabware) equipped with a Ti-sapphire laser (Mai-Tai 'Deepsee', Spectraphysics; wavelength, 920 nm) and a 16x, 0.8 NA water immersion objective (Nikon) at a zoom of 1.6x. The microscope was controlled by Scanbox (Neurolabware) running on MATLAB (Mathworks). Images were acquired at a frame rate of 15.5 (RF mapping) or 31 Hz (occlusion experiments). For all behavioral training and imaging sessions, mice were head restrained and free to run on a 3D-printed running wheel

(8.5 cm diameter). Locomotion speed was measured with a rotary encoder and processed using Arduino and MATLAB.

All visual stimuli for two-photon imaging were presented on a 24-inch gamma-corrected non-occluded HD LED monitor (1920x1080 pixels) using OpenGL and Psychophysics Toolbox 3 running on MATLAB. The monitor was positioned centered at 15 cm away from the mouse, yielding a field-of-view of 120° x 90°.

For RF mapping, we used eccentricity-corrected sparse noise stimuli with squares of 7.5 visual degrees at different locations covering the entire screen. Each trial consisted of 3 black and 3 white randomly positioned squares on a gray background. Stimuli were presented for 0.5 s with a gray screen ITI of 1.5 s.

For occlusion experiments, we presented 6 natural images that were part of a dataset used for similar experiments before[23,27]. The images were presented in gray scale and were corrected for overall luminance across images. We presented images non-occluded or partially occluded using a gray occluder positioned on the top left quadrant of the screen. The occluder had identical luminance to the ITI gray screen. Non-occluded and occluded images (50% of trials each) were both presented 20 times for 1 s with a random ITI of 4-6 s.For surround suppression experiments, we presented circular static black and white gratings (0.05 cycles per degree, 90% contrast) of 4 orientations (0°, 45°, 90° and 135°) and 6 sizes (diameter of 23°, 35°, 46°, 68°, 92° and 123°) at the location of the screen that represented the aggregate RF location of the entire imaging field-of-view. Stimuli were presented 10x each for 1 s with a random ITI of 4-6 s.

For annulus experiments, we presented circular static black and white rings (spatial frequency: 0.05 cycles per degree, 90% contrast) of 4 orientations (0°, 45°, 90° and 135°) and 6 sizes (inner diameter of 23°, 35°, 46°, 68°, 92° and 123°) at the location of the screen that represented the aggregate RF location of the entire imaging field-of-view. The ring width was 4 cm. Stimuli were presented 10x each for 1 s with a random ITI of 4-6 s.

To estimate the influence of cRF and ecRF information on familiar and novel image responding PyCS, we presented 4000 images from 720 classes taken from the THINGS database[70]. Images were shown for 0.5 s, followed by a delay of 0.5 s where mice viewed a gray screen.

**Wide-field imaging**

Wide-field imaging was performed as described as before[71]. Briefly, we imaged the entire 3 mm window using a wide-field fluorescence macroscope (Axio Zoom.V16 Zeiss/Caenotec-Prof. Ralf Schnabel). The mouse was positioned in the center of the LCD screen. Images were captured at 20 Hz by a high-speed sCMOS camera (pco.edge 5.5) and recorded using the Encephalos software package (Caenotec-Prof. Ralf Schnabel).

All visual stimuli used for wide-field imaging were created and presented using COGENT (developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience) running in MATLAB as described before[71]. In short, mice were positioned 14 cm away from a 122 × 68 cm LCD screen (Iiyama LE5564S-B1), resulting in a field-of-view of 144° x 86°. The screen had a resolution of 1920 × 1280 pixels running at refresh rate of 60 Hz. Stimuli used for pRF mapping were as described before[71] and consisted of eccentricity-corrected angled bars (at 0°, 45°, 90° or 135°, 10° in diameter) of a checkerboard pattern on a gray (20 cd/m$^2$) background. The total number of 58 stimuli were presented 10 times each for 0.5 s followed by a gray (20 cd/m$^2$) ITI of 3.6 s. Natural stimuli used for wide-field occlusion experiments were similar as for two-photon experiments (described above). Both non-occluded and occluded images were presented 20 times for 1 s with an ITI of 4-6 s.

**Behavioral training and analysis**

For behavioral training in the reward-based paradigm, water-deprived mice were trained to detect 4 out of 6 non-occluded images to make them familiar. The remaining 2 novel/untrained images were not shown during training. Mice were free to run on the wheel while images were presented for 2 s with a random inter-trial interval of 7-10 s. In the initial passive stage of the training, mice were rewarded with a 5 µl water drop 1 s after stimulus offset. When mice started to lick before the reward delivery, training became active, and a water reward was delivered only when the mice licked in the time window 1-2 s after stimulus offset. Imaging was performed when mice consistently managed to obtain a reward on >70% of trials (generally between 20-30

sessions of 100-200 trials each, ~3750 trials). Mice continued to be trained in between imaging sessions until the experiments were completed.

The hit rate was defined as:

$$Hit\ rate = \frac{N_{hits}}{N_{hits} + N_{misses}}\ x\ 100\% \qquad 1.$$

For behavioral training in the passive paradigm, mice were presented with 4 out of the 6 non-occluded images. The remaining 2 novel/untrained images were not shown during training. No lickspout was present and no rewards were delivered. Each mouse went through a 3-day training regime with 2 training sessions per day (morning and afternoon). Stimuli were presented for 1 s each with a random ITI of 4-6 s. Mice received a total of 3000 presentations per stimulus.

**Wide-field imaging analysis**

Captured images (1600x1600 pixels, 20 Hz, 50 ms exposure) were down sampled to 800x800 pixels. Images were then smoothed using a Gaussian filter with a standard deviation of 2 pixels and stored in a trial based format (400 x 400 x time x number of trials). Analysis of pRF mapping and field-sign was performed as described in detail before[71].

For analysis of occlusion experiments, images were semi-automatically registered using intensity-based rotation and translation to the pRF mapping session[71]. Using the pRF map, we extracted signals for the retinotopic region of V1 corresponding to the occluded quadrant of the screen (with a margin of 7 visual degrees to limit contamination from nearby non-occluded regions of V1). We averaged ΔF/F signals for all pixels in the occluded region and calculated the baseline-corrected (average activity in the 0.2 s before stimulus onset) mean traces across trials and conditions.

**Two-photon imaging analysis**

We used the SpecSeg toolbox for preprocessing as described in detail before[72]. In short, we performed rigid motion correction using NoRMCorre[73] followed by automated region-of-interest (ROI) selection based on cross-spectral power across pixels. After manual refinement, raw ROI signals were extracted and corrected for neuropil by subtracting the average pixel values in an area surrounding each ROI multiplied by 0.7. ΔF/F traces were made by subtracting a moving baseline (10th percentile over 5000 frame windows) and dividing by a linear fit over that baseline.

For RF analysis, we performed reverse correlation on the responses to the sparse noise stimuli to create a 2D response matrix for both the ON (white squares) and OFF (black squares) stimuli. We then fit a circular 2D Gaussian to the response matrix to estimate the RF center and size (full width at half maximum). The quality of the fit was estimated using $r^2$ and the Signal-to-Noise Ration (SNR).

The SNR was defined as:

$$SNR = \frac{Gauss_{peak} - Gauss_{baseline}}{\sigma_{background}} \qquad 2.$$

Where $Gauss_{peak}$ is the maximum of the Gaussian fit, $Gauss_{baseline}$ is the offset parameter and $\sigma_{background}$ is the standard deviation of the background data points (>2 standard deviations away from the center of the Gaussian).

We included neurons if either their ON or their OFF RF met the following criteria: (i) their RF was reliable ($r^2 > 0.33$ and SNR > 4) and (ii) the edge (based on the FWHM) of their RF was at least 2 visual degrees away from the bottom and right-side edges of the occluder. For display purposes (Fig. 2), we used the ON RF for each neuron whose RF met criteria for both the ON and OFF RF map. For neurons recorded in naive and expert mice, we used RFs obtained from sparse noise stimuli shown directly before the occlusion stimuli. For the task data, we used RFs obtained from the expert RF mapping session by matching neurons between the expert and task dataset.

For 2P response strength analyses, we first z-scored ΔF/F signals for each neuron. To compare mean response traces for individual neurons to occlusion stimuli (Fig. 2C), we then averaged across repetitions per image followed by stimulus type (NO and O). To compute single cell average response strengths, we took the baseline-corrected (average 1 s before stimulus onset) mean response between t = 0.2-1 s after stimulus onset.

Average responses to surround suppression and annulus stimuli were computed as described above. We then normalized tuning curves for each neuron to their maximum responses and removed neurons with extreme minimum response values (<-20). We calculated the surround suppression index (SSI) as follows:

$$SSI = \frac{R_{max} - R_{large}}{R_{max}} \quad 3.$$

Where $R_{max}$ is the maximum response to any of the sizes and $R_{large}$ is the response to the largest presented size. We calculated the annulus index (AI) as follows:

$$AI = R_{small} - R_{large} \quad 4.$$

To estimate the relationship between population responses to non-occluded and occluded stimuli across different conditions (Naive, Expert, Task), we calculated the correlation coefficient (R) and the separation index (SI) between non-occluded and occluded responses. For R, we first computed single cell responses to non-occluded and occluded stimuli averaged across images. We then calculated the correlation coefficient for the entire population of neurons by correlating the non-occluded and occluded responses (x-axis and y-axis respectively in Fig. 2F). We calculated the SI as follows. First, we divided the 2D scatter plot (Fig. 2F) into four quadrants (2 x 2). We independently defined the borders between the quadrants for the x-axis and y-axis as the mean + 0.75*STD for the non-occluded and occluded responses, respectively. Second, we counted the number of neurons whose response profile was located in the top right quadrant. To estimate

the degree to which non-occluded and occluded responses were separated, we compared the neuron count in the upper right quadrant of the real distribution to the same count in a non-correlated distribution with independent sampling for the x and y-coordinate. For this, we ran 10,000 repetitions of the count while each time randomly shuffling the non-occluded and occluded responses (creating random pairs of x and y values), resulting in a distribution of permutation counts. We calculated the separation index (SI) as follows:

$$SI = 1 - \frac{Count_{real}}{MeanCount_{perm}} \quad 5.$$

Where $Count_{real}$ is the upper right quadrant count from the real data and $MeanCount_{perm}$ is the distribution average of upper right quadrant counts from the permutation data. Values <-1 were clipped at -1. An SI value of -1 represents a separation lower than the chance distribution (i.e. anti-separation or correlation), 0 represents a lack of separation and 1 represents a separation higher than the chance distribution.

To correlate the strength of the response to the occluded images with the distance of RF edges to the occluder edge for each neuron, we took either the ON or the OFF RF of the neuron depending on which matched the criteria described above. If both the ON and OFF RFs matched the criteria, we took the ON RF. The distance of the RF edge to the occluder was determined as the shortest of the two possible occluder edges (horizontal or vertical).

We calculated response onset latency for each neuron as described before[32,74] by fitting the sum of an exponentially modulated Gaussian and cumulative Gaussian to the activity time courses for non-occluded and occluded conditions during the task. We defined the onset latency as the time point at which the fitted curve reached 33% of its maximum. We only included neurons with a high-quality fit ($r^2 > 0.5$) and an average response (t = 0.2-1 s) greater than 0.

To estimate the contribution of locomotion to responses and plasticity we used two independent procedures. First, we performed linear regression for each neuron by using locomotion speed as a covariate and neuronal activity as a response variable. We then used the residuals of the model

for all subsequent analyses. Second, we repeated the analyses after removing all trials in which the mouse had an average running speed > 2 cm/s during the stimulus window (t = 0.2-1 s).

We defined stimulus selectivity across the 4 different natural images by calculating lifetime sparseness for each neuron as follows:

$$Stimulus\ selectivity = 1 - \frac{(\sum R_i)^2}{N \cdot \sum {R_i}^2} \qquad 6.$$

Where $R_i$ represents the neuron's response to the $i$-th stimulus, and $N$ is the total number of stimuli. We only included responsive neurons (response > 0.3 for that condition).

We calculated selectivity for novel stimuli compared to familiar stimuli in neurons that are responsive to novel stimuli (response > 0.5). For each neuron, we used responses to image pairs in which one image was novel and the other was familiar. For each image pair, we computed a selectivity index as follows:

$$Novel\ selectivity = \frac{R_{novel} - R_{familiar}}{R_{novel} + R_{familiar}} \qquad 7.$$

where $R_{novel}$ and $R_{familiar}$ represent the neuron's response to the novel and familiar image in the pair, respectively. Only pairs for which both responses were greater than zero were included.

We matched neurons across chronically recorded sessions using the chronic matching module of SpecSeg[72]. To correlate the change in response strength between naive and expert conditions to the selectivity to the non-occluded stimulus for individual neurons, we computed the correlation coefficient using the population of neurons that were found back in both sessions.

To estimate the rate of adaptation to different stimuli within individual sessions, we first calculated the average response per trial to each stimulus. We then averaged across stimulus type (familiar, novel, NO and O), resulting in response curves representing the change in response across the 20 trials per stimulus type. We then slightly smoothed these curves with a 1-D Gaussian kernel (SD = 0.5 trials). For the adaptation analysis, we compared the average response

for the first 5 trials in the session to the last 5 trials for each neuron and stimulus type. We only included neurons with a response > 0 averaged across all 20 trials.

We calculated the activity ratio based on reverse correlations (Fig. S9) as follows. For each neuron, we computed the reverse correlation across 3,600 natural images, as the weighted sum of the response to each image (averaged in the time window 0-500ms after the stimulus onset, and zero-centered) and the image structure computed using Shannon's entropy. Image structure, i.e. the entropy, was computed locally in bins of non-overlapping 30×30 pixels, using *skimage*[75]. This resulted in a single image for each neuron, corresponding to their profile of spatial activation. Then, population images were obtained by averaging the individual images of all neurons, separately for each condition. Finally, for the population images (Fig. S9A) we computed the ratio of activity elicited by image parts outside the classical RF (cRF) and the parts inside the cRF. The cRF region was defined at the population level, by summing up the RFs of all the neurons, thresholding the obtained result at the 95$^{th}$ percentile. The SEM of the ratios was computed by bootstrap resampling (1,000 iterations), while statistical differences were tested for significance using a permutation test (1,000 iterations) and corrected for multiple comparisons with Bonferroni.

**Decoding analysis**

For population decoding of mouse WF and 2P data and monkey MUA data, we performed LDA using the 'fitcdiscr' function (default parameters) in MATLAB on the neuronal responses to the natural images. We used individual neurons or pixels as predictors for 2P data and WF data respectively. For trial decoding, responses were calculated by taking the baseline-corrected mean ΔF/F from 0.2 to 1 s after stimulus onset for each trial. Decoding of human fMRI data was adapted from Smith & Muckli (2010)[25]. Decoding over time for monkey MUA data was adapted from Papale *et al.* (2023)[27]. For decoding over time using mouse 2P data, we took the baseline-corrected ΔF/F for each frame and each neuron. On each decoding run, we trained an LDA decoder by randomly selecting 50% of the trials and then testing the decoder on the remaining 50% of trials. We repeated this procedure 200 times to account for variability between trials. We

built two classifiers, one on non-occluded trials (NO model) and one on occluded trials (O model). We then tested each classifier on both conditions to yield four accuracy scores.

Statistical significance for decoding in individual mice (WF decoding), individual frames (i.e. over time), or the overall population was calculated using a permutation test as described before[27]. We first built a null distribution by repeating the decoding procedure 1000 times with shuffled labels. We then selected the upper tail (highest 30%) of the null distribution, and we fitted a Pareto function to the tail to estimate extreme values[76]. The p-value was identified as the percentile corresponding to the 'real' accuracy value in the null-distribution. To compare the decoding accuracy for O-O between conditions (Naive vs Expert), we used the residuals after regressing out locomotion speed to compute the 200 decoding runs for both conditions and calculated the p-value as described above.

**Histology and confocal microscopy**

To examine GCaMP expression, mice were perfused with 15 ml of ice-cold PBS followed by 30 ml of 4% paraformaldehyde (PFA). The brains were extracted and post-fixed for 2 h in 4% PFA at 4°C. Brains were cut in 50 µm slices and imaged on an Axioscan.Z1 slide scanner (ZEISS, Germany) at 20X (0.63 NA).

**Statistical analyses**

All statistical details for each experiment are shown in figures and figure legends and in the statistical table. Unless specified otherwise, statistical analyses were performed using LMEM ('fitlme' function in MATLAB). For LMEM, we considered the response parameter (e.g. visual response magnitude) as the dependent variable, mouse as random effect and the condition (e.g. Naive vs Expert) as the fixed effect. LMEM accounts for the hierarchical correlation structure in the data (i.e. neurons and mice). When the model resulted in a Hessian error (i.e. random effect could not be estimated), we iteratively removed outlier neurons (i.e. neurons with residuals <-

2.5 or >2.5 z-score) before refitting the model. We performed a one-way ANOVA on the LMEM followed by a post hoc coefTest using Tukey's HSD to correct for multiple comparisons.

## Literature


1. Zmarz, P. & Keller, G. B. Mismatch Receptive Fields in Mouse Visual Cortex. *Neuron* **92**, 766–772 (2016).
2. Lamme, V. A. The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* **15**, 1605–1615 (1995).
3. Homann, J., Koay, S. A., Chen, K. S., Tank, D. W. & Berry, M. J. Novel stimuli evoke excess activity in the mouse primary visual cortex. *Proc Natl Acad Sci U S A* **119**, e2108882119 (2022).
4. Furutachi, S., Franklin, A. D., Aldea, A. M., Mrsic-Flogel, T. D. & Hofer, S. B. Cooperative thalamocortical circuit mechanism for sensory prediction errors. *Nature* **633**, 398–406 (2024).
5. Hamm, J. P. & Yuste, R. Somatostatin Interneurons Control a Key Component of Mismatch Negativity in Mouse Visual Cortex. *Cell Reports* **16**, 597–604 (2016).
6. Mendoza-Halliday, D., Xu, H., Azevedo, F. A. C. & Desimone, R. Dissociable neuronal substrates of visual feature attention and working memory. *Neuron* **112**, 850-863.e6 (2024).
7. Keller, G. B., Bonhoeffer, T. & Hübener, M. Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron* **74**, 809–815 (2012).
8. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* **36**, 181–204 (2013).
9. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* **2**, 79–87 (1999).
10. Friston, K. Functional integration and inference in the brain. *Prog Neurobiol* **68**, 113–143 (2002).
11. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron* **100**, 424–435 (2018).
12. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
13. Shipp, S. Neural Elements for Predictive Coding. *Front Psychol* **7**, 1792 (2016).

14. Jordan, R. & Keller, G. B. Opposing Influence of Top-down and Bottom-up Input on Excitatory Layer 2/3 Neurons in Mouse Primary Visual Cortex. *Neuron* **108**, 1194-1206.e5 (2020).

15. Liu, Y. *et al.* Organization of corticocortical and thalamocortical top-down inputs in the primary visual cortex. *Nat Commun* **15**, 4495 (2024).

16. Larkum, M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends in Neurosciences* **36**, 141–151 (2013).

17. Allman, J., Miezin, F. & McGuinness, E. Stimulus Specific Responses from Beyond the Classical Receptive Field: Neurophysiological Mechanisms for Local-Global Comparisons in Visual Neurons. *Annual Review of Neuroscience* **8**, 407–430 (1985).

18. Niell, C. M. & Stryker, M. P. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron* **65**, 472–479 (2010).

19. Iurilli, G. *et al.* Sound-Driven Synaptic Inhibition in Primary Visual Cortex. *Neuron* **73**, 814–828 (2012).

20. Meijer, G. T., Montijn, J. S., Pennartz, C. M. A. & Lansink, C. S. Audiovisual Modulation in Mouse Primary Visual Cortex Depends on Cross-Modal Stimulus Configuration and Congruency. *J Neurosci* **37**, 8783–8796 (2017).

21. Muckli, L. & Petro, L. S. Network interactions: non-geniculate input to V1. *Current Opinion in Neurobiology* **23**, 195–201 (2013).

22. Muckli, L. *et al.* Contextual Feedback to Superficial Layers of V1. *Current Biology* **25**, 2690–2695 (2015).

23. Morgan, A. T., Petro, L. S. & Muckli, L. Scene Representations Conveyed by Cortical Feedback to Early Visual Cortex Can Be Described by Line Drawings. *J. Neurosci.* **39**, 9410–9423 (2019).

24. Revina, Y., Petro, L. S. & Muckli, L. Cortical feedback signals generalise across different spatial frequencies of feedforward inputs. *NeuroImage* **180**, 280–290 (2018).

25. Smith, F. W. & Muckli, L. Nonstimulated early visual areas carry information about surrounding context. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20099–20103 (2010).

26. Petro, L. S., Smith, F. W., Abbatecola, C. & Muckli, L. The Spatial Precision of Contextual Feedback Signals in Human V1. *Biology* **12**, 1022 (2023).

27. Papale, P. *et al.* The representation of occluded image regions in area V1 of monkeys and humans. *Current Biology* **33**, 3865-3871.e3 (2023).

28. Sawtell, N. B. *et al.* NMDA Receptor-Dependent Ocular Dominance Plasticity in Adult Visual Cortex. *Neuron* **38**, 977–985 (2003).

29. Dumoulin, S. O. & Wandell, B. A. Population receptive field estimates in human visual cortex. *NeuroImage* **39**, 647–660 (2008).

30. Erisken, S. *et al.* Effects of Locomotion Extend throughout the Mouse Early Visual System. *Current Biology* **24**, 2899–2907 (2014).

31. Keller, A. J., Roth, M. M. & Scanziani, M. Feedback generates a second receptive field in neurons of the visual cortex. *Nature* **582**, 545–549 (2020).

32. Kirchberger, L., Mukherjee, S., Self, M. W. & Roelfsema, P. R. Contextual drive of neuronal responses in mouse V1 in the absence of feedforward input. *Sci Adv* **9**, eadd2498 (2023).

33. Makino, H. & Komiyama, T. Learning enhances the relative impact of top-down processing in the visual cortex. *Nat Neurosci* **18**, 1116–1122 (2015).

34. Abbott, L. F., Varela, J. A., Sen, K. & Nelson, S. B. Synaptic depression and cortical gain control. *Science* **275**, 220–224 (1997).

35. Tsodyks, M. V. & Markram, H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci U S A* **94**, 719–723 (1997).

36. Grubb, M. S. & Burrone, J. Activity-dependent relocation of the axon initial segment fine-tunes neuronal excitability. *Nature* **465**, 1070–1074 (2010).

37. Evans, M. D., Dumitrescu, A. S., Kruijssen, D. L. H., Taylor, S. E. & Grubb, M. S. Rapid Modulation of Axon Initial Segment Length Influences Repetitive Spike Firing. *Cell Rep* **13**, 1233–1245 (2015).

38. Desai, N. S., Rutherford, L. C. & Turrigiano, G. G. Plasticity in the intrinsic excitability of cortical pyramidal neurons. *Nat Neurosci* **2**, 515–520 (1999).

39. Ibata, K., Sun, Q. & Turrigiano, G. G. Rapid synaptic scaling induced by changes in postsynaptic firing. *Neuron* **57**, 819–826 (2008).

40. Xue, M., Atallah, B. V. & Scanziani, M. Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature* **511**, 596–600 (2014).

41. Schulz, A., Miehl, C., Berry, M. J. & Gjorgjieva, J. The generation of cortical novelty responses through inhibitory plasticity. *Elife* **10**, e65309 (2021).

42. Kok, P., Jehee, J. F. M. & de Lange, F. P. Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron* **75**, 265–270 (2012).

43. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb Cortex* **16**, 1631–1644 (2006).

44. Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences* **10**, 14–23 (2006).

45. Kumar, S., Kaposvari, P. & Vogels, R. Encoding of Predictable and Unpredictable Stimuli by Inferior Temporal Cortical Neurons. *Journal of Cognitive Neuroscience* **29**, 1445–1454 (2017).

46. De Lange, F. P., Heilbron, M. & Kok, P. How Do Expectations Shape Perception? *Trends in Cognitive Sciences* **22**, 764–779 (2018).

47. Richter, D., Heilbron, M. & De Lange, F. P. Dampened sensory representations for expected input across the ventral visual stream. *Oxford Open Neuroscience* **1**, kvac013 (2022).

48. Blank, H. & Davis, M. H. Prediction Errors but Not Sharpened Signals Simulate Multivoxel fMRI Patterns during Speech Perception. *PLoS Biol* **14**, e1002577 (2016).

49. Han, B., Mostert, P. & De Lange, F. P. Predictable tones elicit stimulus-specific suppression of evoked activity in auditory cortex. *NeuroImage* **200**, 242–249 (2019).

50. Wright, W. J., Hedrick, N. G. & Komiyama, T. Distinct synaptic plasticity rules operate across dendritic compartments in vivo during learning. *Science* **388**, 322–328 (2025).

51. Gambino, F. *et al.* Sensory-evoked LTP driven by dendritic plateau potentials in vivo. *Nature* **515**, 116–119 (2014).

52. Bastos, G. *et al.* Top-down input modulates visual context processing through an interneuron-specific circuit. *Cell Rep* **42**, 113133 (2023).

53. Williams, L. E. & Holtmaat, A. Higher-Order Thalamocortical Inputs Gate Synaptic Long-Term Potentiation via Disinhibition. *Neuron* **101**, 91-102.e4 (2019).

54. van Versendaal, D. & Levelt, C. N. Inhibitory interneurons in visual cortical plasticity. *Cellular and Molecular Life Sciences* **73**, 3677–3691 (2016).

55. Hofer, S. B. *et al.* Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat Neurosci* **14**, 1045–1052 (2011).

56. Shin, H. *et al.* Recurrent pattern completion drives the neocortical representation of sensory inference. *Nat Neurosci* **28**, 2319–2329 (2025).

57. Major, G., Larkum, M. E. & Schiller, J. Active properties of neocortical pyramidal neuron dendrites. *Annu Rev Neurosci* **36**, 1–24 (2013).

58. Larkum, M. E., Zhu, J. J. & Sakmann, B. A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* **398**, 338–341 (1999).

59. Larkum, M. E., Waters, J., Sakmann, B. & Helmchen, F. Dendritic spikes in apical dendrites of neocortical layer 2/3 pyramidal neurons. *J Neurosci* **27**, 8999–9008 (2007).

60. Marvan, T. & Phillips, W. A. Cellular mechanisms of cooperative context-sensitive predictive inference. *Curr Res Neurobiol* **6**, 100129 (2024).

61. Marvan, T., Polák, M., Bachmann, T. & Phillips, W. A. Apical amplification—a cellular mechanism of conscious perception? *Neurosci Conscious* **2021**, niab036 (2021).

62. Dias, R. F. *et al.* Visual experience reduces the spatial redundancy between cortical feedback inputs and primary visual cortex neurons. *Neuron* **112**, 3329-3342.e7 (2024).

63. Rajan, R. *et al.* Visual experience exerts an instructive role on cortical feedback inputs to the primary visual cortex. *Curr Biol* S0960-9822(26)00070–9 (2026) doi:10.1016/j.cub.2026.01.031.

64. Summerfield, C., Wyart, V., Johnen, V. M. & De Gardelle, V. Human Scalp Electroencephalography Reveals that Repetition Suppression Varies with Expectation. *Front. Hum. Neurosci.* **5**, (2011).

65. Meyer, T. & Olson, C. R. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19401–19406 (2011).

66. Todorovic, A. & De Lange, F. P. Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields. *J. Neurosci.* **32**, 13389–13395 (2012).

67. Wacongne, C. *et al.* Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20754–20759 (2011).

68. Von Helmholtz, H. *Handbuch der physiologischen Optik*. (1867).

69. Keller, G. B. & Sterzer, P. Predictive Processing: A Circuit Approach to Psychosis. *Annu Rev Neurosci* **47**, 85–101 (2024).

70. Hebart, M. N. *et al.* THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PLOS ONE* **14**, e0223792 (2019).

71. Van Beest, E. H. *et al.* Mouse visual cortex contains a region of enhanced spatial resolution. *Nat Commun* **12**, 4029 (2021).

72. De Kraker, L. *et al.* SpecSeg is a versatile toolbox that segments neurons and neurites in chronic calcium imaging datasets based on low-frequency cross-spectral power. *Cell Reports Methods* **2**, 100299 (2022).

73. Pnevmatikakis, E. A. & Giovannucci, A. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods* **291**, 83–94 (2017).

74. 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).

75. Walt, S. van der *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).

76. Knijnenburg, T. A., Wessels, L. F. A., Reinders, M. J. T. & Shmulevich, I. Fewer permutations, more accurate *P* -values. *Bioinformatics* **25**, i161–i168 (2009).

77. Syeda, A. *et al.* Facemap: a framework for modeling neural activity based on orofacial tracking. *Nat Neurosci* **27**, 187–195 (2024).

## Key resource table

| Reagent or resource | Source | Identifier |
|---|---|---|
| **Mouse strains** | | |
| G35-3-Cre | Tonegawa lab (MIT) | MGI:4455188 |
| TIT2L-GCaMP6f (Ai148D) | Jackson Laboratories | 030328 |

# Statistical table

| Figure | Comparison | Parameter | Test | # mice | N represents | # neurons | p-val | Sign |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| **2E** | Naive vs Expert | NO response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 884 | <0.001 | *** |
| | Naïve vs Task | NO response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 300 | <0.001 | *** |
| | Expert vs Task | NO response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 300 | 0.9999 | ns |
| | Naive vs Expert | O response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 884 | <0.001 | *** |
| | Naïve vs Task | O response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 300 | <0.001 | *** |
| | Expert vs Task | O response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 300 | <0.001 | *** |
| | | | | | | | | |
| **2G** | NO vs O | Response onset | LMEM-Anova | 6 vs 6 | Neurons | 584 vs 528 | <0.001 | *** |
| | | | | | | | | |
| **2H** | Naive vs Expert | NO selectivity | LMEM-Anova | 6 vs 6 | Neurons | 328 vs 223 | <0.001 | *** |
| | Naïve vs Task | NO selectivity | LMEM-Anova | 6 vs 6 | Neurons | 328 vs 79 | <0.001 | *** |
| | Expert vs Task | NO selectivity | LMEM-Anova | 6 vs 6 | Neurons | 223 vs 79 | <0.001 | *** |
| | Naive vs Expert | O selectivity | LMEM-Anova | 6 vs 6 | Neurons | 79 vs 182 | 0.9999 | ns |
| | Naïve vs Task | O selectivity | LMEM-Anova | 6 vs 6 | Neurons | 79 vs 79 | 0.9999 | ns |
| | Expert vs Task | O selectivity | LMEM-Anova | 6 vs 6 | Neurons | 182 vs 79 | 0.9999 | ns |
| | | | | | | | | |
| **3C** | N vs PE | NO response | LMEM-Anova | 7 vs 7 | Neurons | 1820 vs 2106 | <0.001 | *** |
| | N vs PE | O response | LMEM-Anova | 7 vs 7 | Neurons | 1820 vs 2106 | <0.001 | *** |
| | | | | | | | | |
| **3F** | NO cells vs O cells (N) | Surround suppression | LMEM-Anova | 7 vs 7 | Neurons | 691 vs 61 | 0.302 | ns |
| | NO cells vs O cells (PE) | Surround suppression | LMEM-Anova | 7 vs 7 | Neurons | 392 vs 100 | <0.001 | *** |
| | | | | | | | | |
| **3I** | NO cells vs O cells (N) | Annulus index | LMEM-Anova | 7 vs 7 | Neurons | 691 vs 61 | <0.001 | *** |
| | NO cells vs O cells (PE) | Annulus index | LMEM-Anova | 7 vs 7 | Neurons | 392 vs 100 | <0.001 | *** |
| | | | | | | | | |
| **3J** | NO-NO vs chance | Decoding accuracy | LMEM-Anova | 7 | Neurons | 2106 | <0.001 | *** |
| | O-O vs chance | Decoding accuracy | LMEM-Anova | 7 | Neurons | 2106 | <0.001 | *** |
| | | | | | | | | |
| **4C** | Naive vs Expert | NO response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 884 | <0.001 | *** |
| | Naive vs Expert | O response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 884 | <0.001 | *** |
| | | | | | | | | |
| **4G** | Naive vs Expert | NO selectivity | LMEM-Anova | 6 vs 6 | Neurons | 172 vs 285 | <0.001 | *** |
| | | | | | | | | |
| **5A** | NO-NO vs chance | Decoding accuracy | Permutation test | 6 | Neurons | 300 | <0.001 | *** |
| | O-O vs chance | Decoding accuracy | Permutation test | 6 | Neurons | 300 | 0.001 | ** |
| | NO-O vs chance | Decoding accuracy | Permutation test | 6 | Neurons | 300 | 0.018 | * |
| | O-NO vs chance | Decoding accuracy | Permutation test | 6 | Neurons | 300 | 0.001 | ** |
| | | | | | | | | |
| **5B** | NO-NO vs chance | Decoding accuracy | Permutation test | 175 | Recording sites | NA | <0.001 | *** |
| | O-O vs chance | Decoding accuracy | Permutation test | 175 | Recording sites | NA | <0.001 | *** |
| | NO-O vs chance | Decoding accuracy | Permutation test | 175 | Recording sites | NA | 0.007 | ** |
| | O-NO vs chance | Decoding accuracy | Permutation test | 175 | Recording sites | NA | <0.001 | *** |
| | NO-NO vs chance | Decoding accuracy | t-test | 6 | Humans | NA | <0.001 | *** |
| | O-O vs chance | Decoding accuracy | t-test | 6 | Humans | NA | 0.0058 | ** |
| | NO-O vs chance | Decoding accuracy | t-test | 6 | Humans | NA | 0.0081 | ** |
| | O-NO vs chance | Decoding accuracy | t-test | 6 | Humans | NA | 0.0038 | ** |
| | | | | | | | | |
| **S2A** | Naive vs Expert | NO response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 841 | <0.001 | *** |
| | Naïve vs Task | NO response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 300 | <0.001 | *** |
| | Expert vs Task | NO response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 300 | 0.9999 | ns |
| | Naive vs Expert | O response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 841 | <0.001 | *** |
| | Naïve vs Task | O response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 300 | <0.001 | *** |
| | Expert vs Task | O response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 300 | <0.001 | *** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **S2B** | Naive vs Expert | NO response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 841 | <0.001 | *** |
| | Naïve vs Task | NO response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 300 | 0.0333 | * |
| | Expert vs Task | NO response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 300 | 0.2348 | ns |
| | Naive vs Expert | O response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 841 | <0.001 | *** |
| | Naïve vs Task | O response | LMEM-Anova | 6 vs 6 | Neurons | 884 vs 300 | <0.001 | *** |
| | Expert vs Task | O response | LMEM-Anova | 6 vs 6 | Neurons | 841 vs 300 | <0.001 | *** |
| | | | | | | | | |
| **S2D** | Position vs zero, NO, X | Pupil position | Signrank | 6 vs 6 | Mice | NA | 0.6875 | ns |
| | Position vs zero, NO, Y | Pupil position | Signrank | 6 vs 6 | Mice | NA | 0.9999 | ns |
| | Position vs zero, O, X | Pupil position | Signrank | 6 vs 6 | Mice | NA | 0.5625 | ns |
| | Position vs zero, O, Y | Pupil position | Signrank | 6 vs 6 | Mice | NA | 0.3125 | ns |
| | | | | | | | | |
| **S9B** | NO fam vs O fam | Activity ratio | Permutation test | 6 vs 6 | Neurons | 277 vs 63 | 0.0060 | ** |
| | NO fam vs NO nov | Activity ratio | Permutation test | 6 vs 6 | Neurons | 277 vs 433 | 0.0060 | ** |
| | NO fam vs O nov | Activity ratio | Permutation test | 6 vs 6 | Neurons | 277 vs 52 | 0.0060 | ** |
| | O fam vs NO nov | Activity ratio | Permutation test | 6 vs 6 | Neurons | 63 vs 433 | 0.0060 | ** |
| | NO nov vs O nov | Activity ratio | Permutation test | 6 vs 6 | Neurons | 433 vs 52 | 0.0060 | ** |
| | O fam vs O nov | Activity ratio | Permutation test | 6 vs 6 | Neurons | 63 vs 52 | 0.1080 | ns |

## Author contributions

| | KS | LdK | PP | LP | JM | MS | ML | PR | LM | CL |
|---|---|---|---|---|---|---|---|---|---|---|
| **Experimental design** | X | | | X | | | | | X | X |
| **Data collection** | X | | | | | | | | | |
| **Data analysis** | X | X | X | | X | X | | | | |
| **Technical assistance** | | | | | | | | | | |
| **Interpretation** | X | | X | | X | | | X | X | X |
| **Writing** | X | | | | | | | | | X |
| **Review/editing** | X | X | | X | X | X | X | X | X | X |
| **Conceptualization** | X | | | | | | | | X | X |
| **Funding acquisition** | | | | X | | | X | | X | X |


## Acknowledgements

We thank all members of the Levelt lab for discussion and support. We thank staff of the animal facility and mechatronics department at the Netherlands Institute for Neuroscience for technical support. We thank Dr. Chris van der Togt, Dr. Enny van Beest, Paul Neering, Maaike van der Aa, Marija Žuvela, Jessie Lof, Augustijn Vrolijk and Anda de Witte for help with experiments and analyses. We thank Dr. William Phillips for the critical reading of the manuscript. We thank Dr.

Fraser Smith for help with the human data. This project received funding from the European Union's Horizon 2020 Research and Innovation Program under grant agreement nos. 785907 (HBP SGA2, CL, PR, ML & LM) and 945539 (HBP SGA3, CL, PR, ML & LM), and from the Biotechnology and Biological Sciences Research Council (BBSRC) "Layer-specific cortical feedback (LM with LSP). The authors declare that they have no competing interests. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

## Supplementary figures

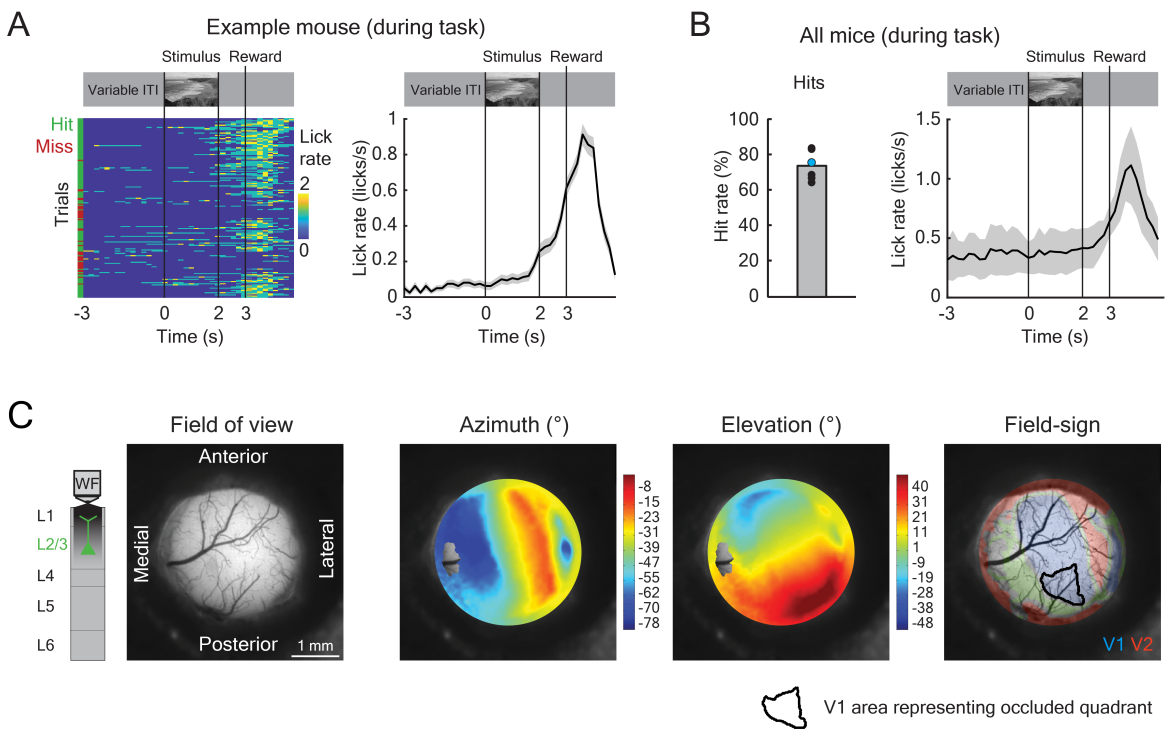


Figure S1. Task behavior and widefield population receptive field maps.

A) Schematic of task structure and licking behavior for an example mouse (blue dot in B). Water restricted mice were required to lick in the interval of 1-2 s after disappearance of a non-occluded natural scene to obtain a water reward. Left: lick rate for individual trials. Right: trial-averaged average lick rate.

B) Left: average hit rate for individual mice during the task session. The blue dots corresponds to the example mouse in A. Right: trial-averaged lick rate (mean ± SEM across 6 mice).

C) Cortical wide-field maps for an example mouse showing from left to right: the imaging field of view, visual azimuth, visual elevation and visual field-sign overlaid on the brain imaged through the glass window. The large blue patch in the field-sign image represents V1 while the red patches represent higher visual areas (V2).

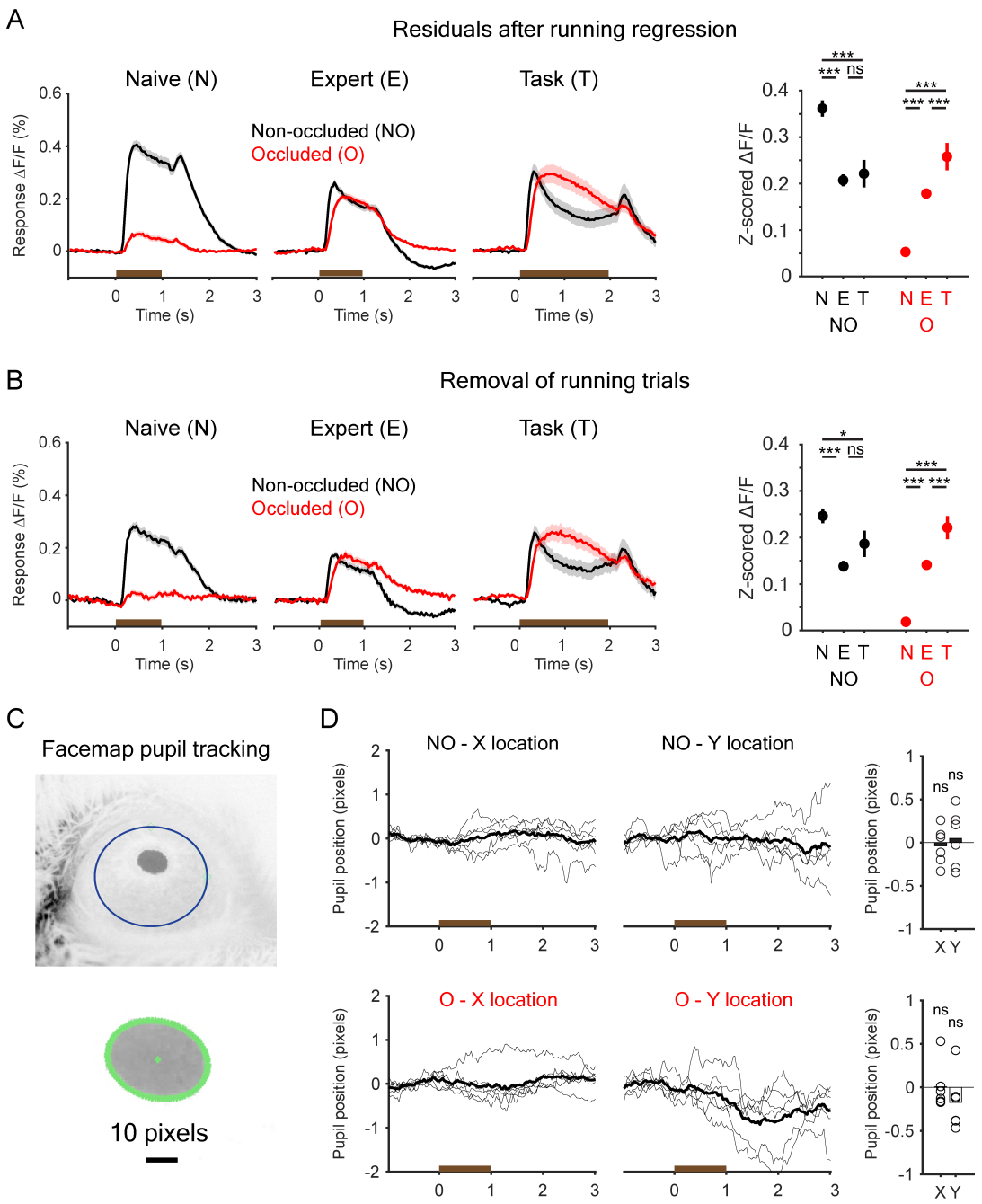


Figure S2. Locomotion and pupil position do not explain experience-dependent plasticity.

A) Left: Schematic and z-scored ΔF/F traces averaged across neurons and stimuli for all conditions after regressing out locomotion speed and performing analyses on the residual data. Brown bars indicate the presence of the visual stimulus. Right: neuron-averaged response strength across stimuli and conditions (mean ± SEM, averaged between t = 0.2-1 s) for neurons after regressing out locomotion speed. N = 841 neurons (Naive), 844 neurons (Expert) and 300 neurons (Task) from 6 mice. LMEM with post hoc Tukey HSD for all comparisons here and below, ***: p<0.001, **: p<0.01, *: p<0.05, ns: not significant.

B) As in A, but for data after removal of locomotion trials (average speed > 2 cm/s).

C) Top: example image of pupil tracking using Facemap[77]. The blue circle represents the eye ROI drawn by the user within which Facemap will track the pupil. Bottom: zoomed in image of the pupil with the tracked pupil ROI (green circle) for an example frame.

D) Traces for pupil position along the X (left) and Y (center) axes of the tracked pupil center for non-occluded (NO, top) and occluded (O, bottom) images. The brown bar indicates stimulus presence. Right panels show the average pupil position per mouse (black circles) during stimulus presentation (t = 0-1 s) relative to baseline (t < 0 s). Signrank test against zero, ns: not significant.

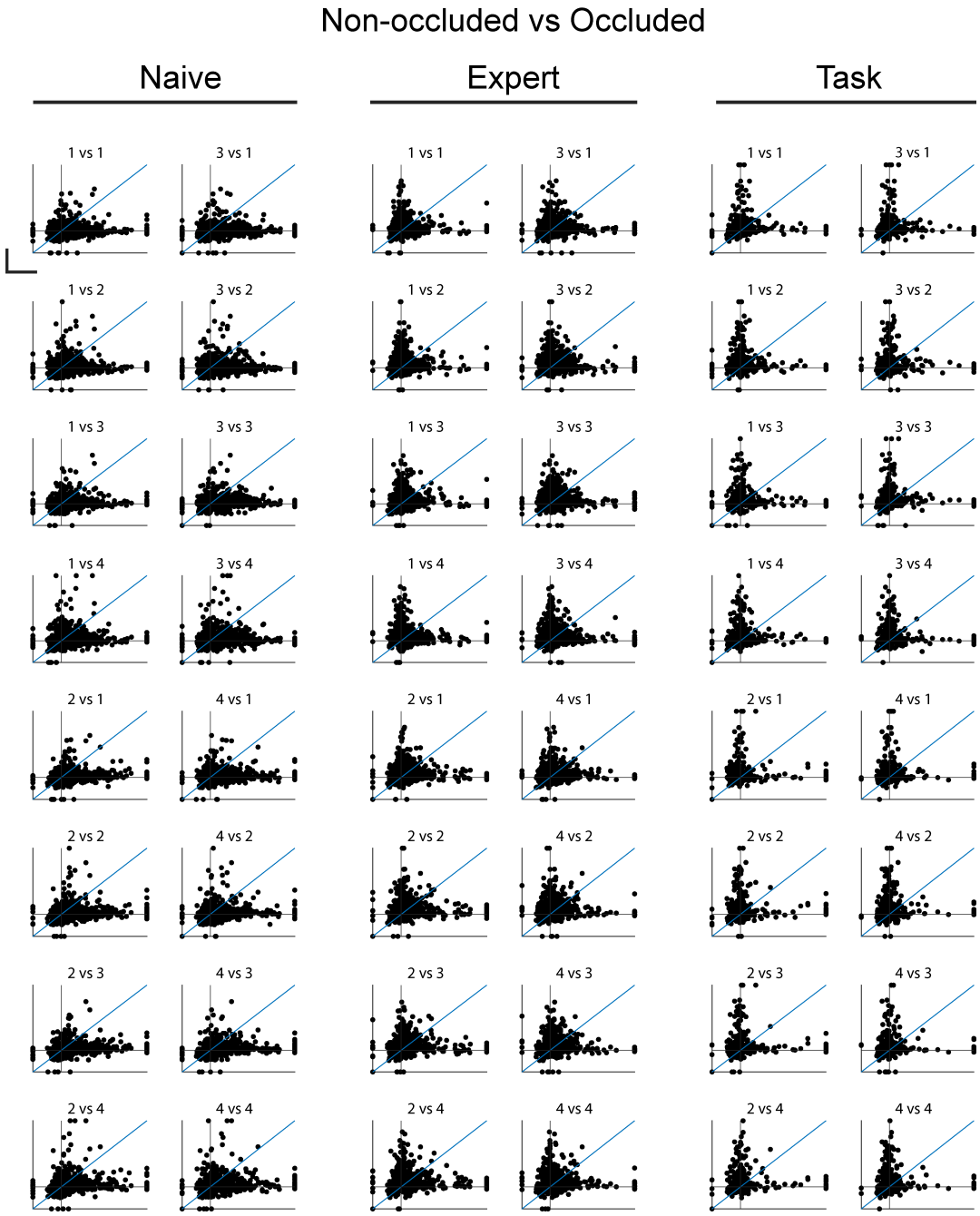


Figure S3. Relationship between responses to non-occluded and occluded images.

Scatter plots showing the relationship between responses to pairs of non-occluded and occluded images. Panel titles indicate the image comparisons made (non-occluded on x-axis vs. occluded on y-axis). Axes are capped at -0.5 and 2.5 z-score. Scale bars, 1 z-scored ΔF/F.

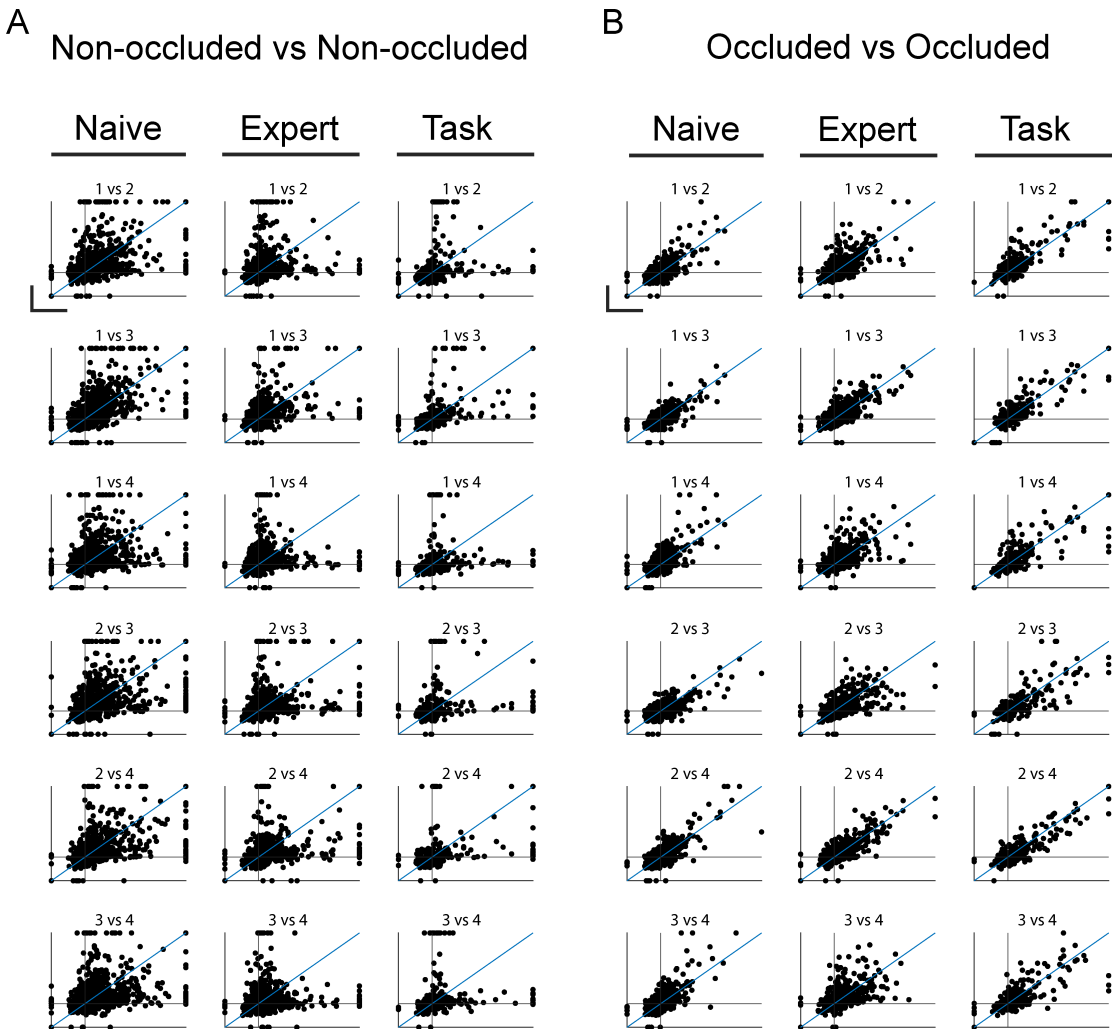


Figure S4. Relationship between responses to non-occluded and occluded images.

A) Scatter plots showing the relationship between responses to pairs of non-occluded images. Panel titles indicate the image comparisons made (non-occluded vs non-occluded). Axes are capped at -0.5 and 2.5 z-score. Scale bars, 1 z-scored ΔF/F.

B) As in A, but for pairs of occluded images.

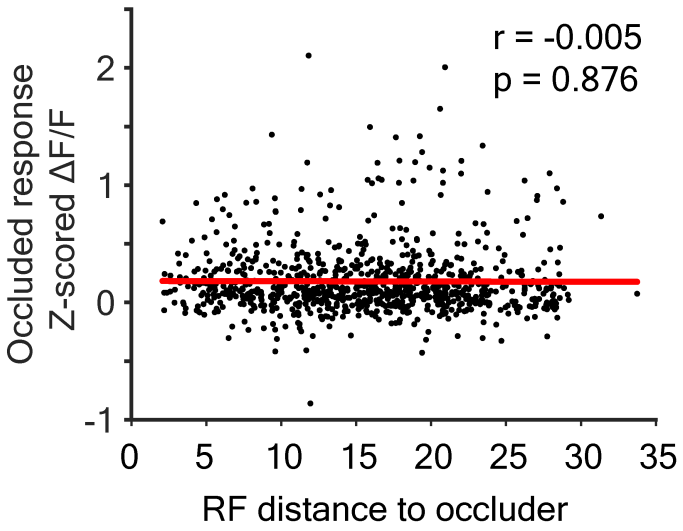


Figure S5. Relationship between occluded response strength and RF distance to the occlude in expert mice. Scatter plot showing the relationship between the distance between RF edge and the border of the occluder and the response to occluded stimuli. The linear regression line is indicated in red.

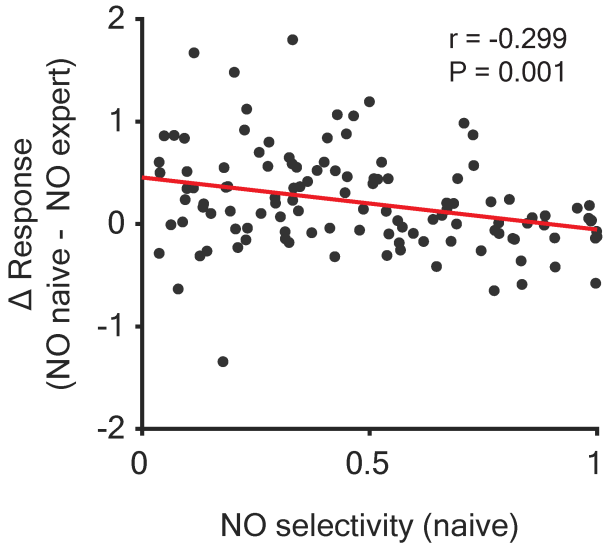


Figure S6. Relationship between non-occluded selectivity and plasticity. Scatter plot showing the relationship between selectivity to non-occluded images and the change in response strength from Naive to Expert conditions (n = 117 PyCs). NO, non-occluded.

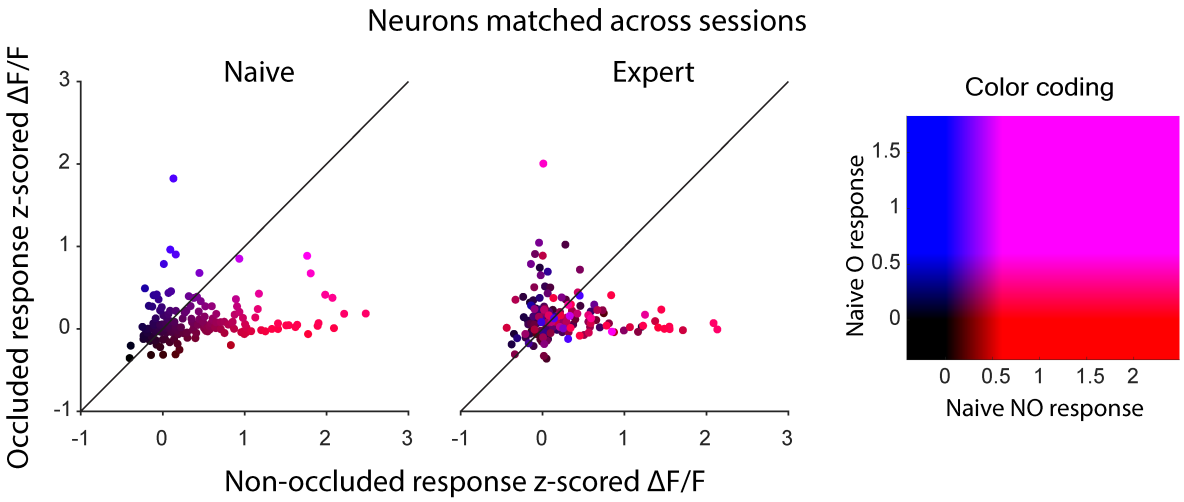


Figure S7. Neurons matched across chronic imaging sessions for Naive vs Expert NO-O responses. Scatter plot showing average z-scored ΔF/F responses (t = 0.2-1 s) to the non-occluded (x-axis) and occluded (y-axis) stimulus of chronically matched cells in naive and expert mice, color coded for response strength in naive mice. N = 166 neurons from 6 mice. NO, non-occluded; O, occluded.

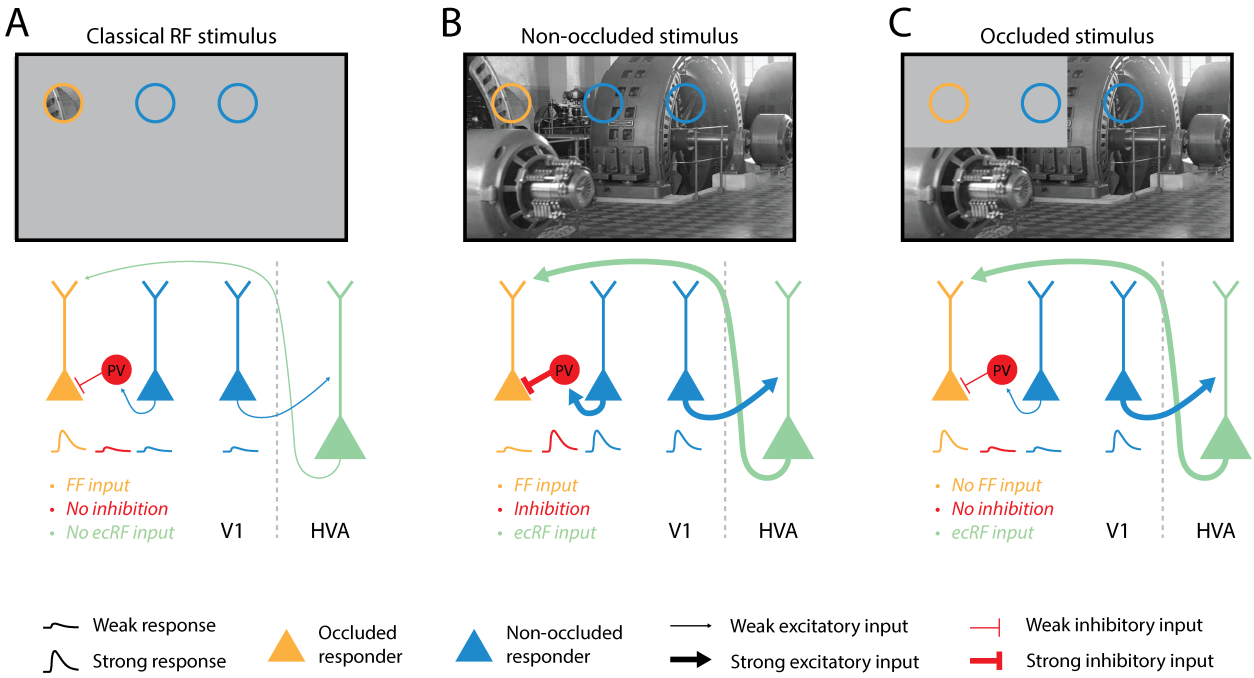


Figure S8. Mechanism underlying selective responses to occluded images

A) A small stimulus in the classical (cRF) drives the left PyC (orange, occluded responder). The PyC receives little inhibition (red) and no extra-classical RF (ecRF) inputs (green).
B) A full visual scene gives input to the orange PyC, but also activates other PyCs (blue, non-occluded responders) that activate local interneurons (red), inhibiting the response of the orange PyC, despite ecRF input.
C) An occluded visual scene, lacking of feedforward input, reduces local inhibition. Combined with ecRF inputs, the orange PyC is activated.

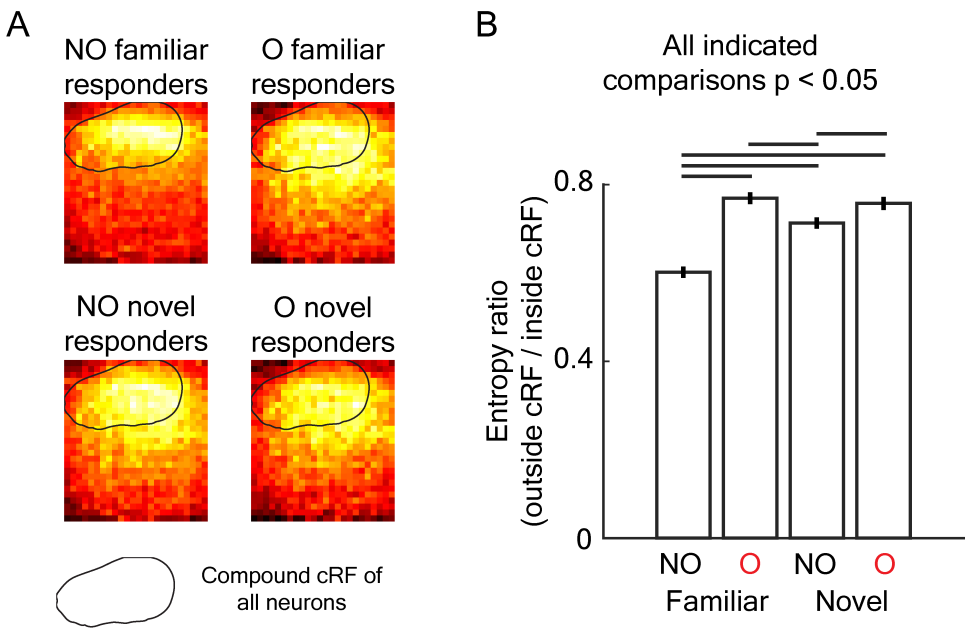


Figure S9. Image structure representations and outside-inside receptive field entropy ratios across L2/3 PyC categories

A) Average image structure representations per neuronal response type. Panels show averaged entropy across all reverse correlated natural images, which captures the complexity of pixel intensities within an image region, for the indicated categories of PyCs (selected based on response > 0.5 for the indicated stimulus). The black outline represents the compound cRF of all neurons.

B) Ratios of outside to inside entropy values per category. Values were computed as the ratio of entropy elicited by image parts outside the cRF and the parts inside the cRF. N = 277 neurons (NO familiar), 63 neurons (O familiar), 433 neurons (NO novel) and 52 neurons (O novel) from 7 mice. All indicated comparisons $p < 0.05$, permutation test with Bonferroni correction.