Unexpected events modulate context signaling in VIP and excitatory cells of the visual cortex

Farzaneh Najafi[1,4 #], Simone Russo[3,5,6], Jérôme Lecoq[1,2 #]

[1] Allen Institute, Mindscope Program, [2] Neural Dynamics Program, [3] Brain and Consciousness Program, Seattle, WA, USA

[4] Georgia Institute of Technology, Biological Sciences, [5] Wallace H Coulter Department of Biomedical Engineering, Atlanta, GA, USA

[6] Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan 20157, Italy

Corresponding authors:

Jerome Lecoq: jeromel@alleninstitute.org

Farzaneh Najafi: fnajafi3@gatech.edu

ABSTRACT

The visual cortex predicts incoming sensory stimuli through internal models of the world. Unexpected stimuli that violate these predictions update internal models and drive adaptation. Cortical inhibitory neurons, particularly VIP (vasoactive intestinal peptide) interneurons, are suggested to play a key role in representing unexpected stimuli, given their robust firing following unexpected omissions of familiar images. Importantly, this response is stimulus non-specific, raising an important question about what information it conveys. Given their unique connectivity with other cell types and brain areas, we hypothesized that during unexpected events, VIP neurons encode contextual information, defined as neuronal activity that is not driven by the stimulus itself. To test this hypothesis, we analyzed the Allen Institute Visual Behavior dataset, in which mice viewed repeated familiar images and unexpected omissions of these images while two-photon calcium imaging data from different cell types and visual areas were recorded. Using dimensionality reduction techniques, we found that omissions trigger contextual signaling in VIP neurons in the primary visual cortex (V1) and the lateral medial (LM) visual area, particularly in the superficial layers. Similarly, contextual coding was enhanced in

excitatory neurons following omissions. This contrasted sharply with the excitatory response to expected images, during which contextual information was substantially suppressed. Our results suggest that unexpected events activate VIP neurons, which subsequently propagate contextual information across the cortical network. This potentially facilitates the integration of context within the cortical network, and leads to updated predictions about our dynamic environment.

INTRODUCTION

The visual cortex is crucial in shaping our internal visual representations of the sensory environment. In the mouse visual cortex, inhibitory neurons are diverse, with major classes including parvalbumin (PV), vasoactive intestinal polypeptide (VIP), and somatostatin (SST) neurons, each with distinct subclasses as evidenced by transcriptomic characterization[1]. VIP cells, presumed to exert a dis-inhibitory effect on pyramidal cells, may play a key role in regulating information flow within the cortical network due to their connectivity with associative cortical areas and neuromodulatory systems[2–5]. VIP neurons exist in superficial (layer 2/3 primarily), and deeper (layer 4) layers of the cortex, suggesting their involvement in both top-down and bottom-up information processing, respectively[6].

Circuit models suggest that unexpected stimuli trigger computations across both excitatory and inhibitory cortical cells to refine sensory representations[7]. Empirical data indicate that unexpected stimuli elicit responses across multiple layers of the cortex, involving inhibitory neurons[8]. A recent study demonstrated that the activity of VIP cells in the visual cortex peaks during unexpected omissions of visual stimuli[9], but the underlying nature of these responses remains unclear. VIP cells may signal the presence of a prediction error signal; alternatively, they may encode contextual changes that co-occur during omissions, including behavioral, external, or internal states.

In the present study, we addressed this question using multi-area two-photon calcium imaging during a visually guided behavior. This task contained both expected events (regular flashes of images) and unexpected events (image omissions). We leveraged population activity decoding, dimensionality reduction and trial shuffling to

examine whether VIP activity during omissions may contain stimulus-specific prediction-error signals, or if it was primarily influenced by context information. Importantly, "context" in our study broadly refers to shared activity across brain areas that are independent of task-relevant features. This may involve behavioral features such as the animal's locomotion, external inputs such as the sensory signals in the animal's environment, and internal representations related to the animal's current state.

We found that most VIP neurons, except those in layer 4 of the primary visual area, encoded context information. This suggests that VIP neurons represent a general 'surprise' signal enriched with contextual details during unexpected events. In contrast, excitatory neurons displayed weaker modulation by context during expected image presentations but were significantly influenced by context during unexpected omissions. These results highlight the role of superficial VIP neurons in updating cortical activity with contextual information in response to unexpected stimuli. This mechanism potentially facilitates the association of contextual signals with unexpected events, thereby driving adaptive responses.

RESULTS

Distinct cortical pathways are thought to carry sensory and prediction signals[10,11], suggesting that cortical areas and cell types might show unique responses to expected vs. unexpected events. To test this hypothesis, we leveraged two-photon imaging data from the Allen Brain Observatory platform[9,12]. In this dataset, mice were trained to perform a go/no-go, image-change detection task[9]. In brief, water-restricted animals were presented with a constant stream of natural images (250 ms) interleaved with gray screens of matched luminance (500 ms). When a change in image identity occurred, mice would receive a water reward if they licked immediately after the image change. Importantly, during imaging sessions in well-trained mice, a pseudo-random 5% of non-change images were replaced with a gray screen (image omissions; Fig. 1A). Since mice had extensive experience with the highly regular timing of image presentations (every ~750 ms), the rare stimulus omissions were unexpected events, as opposed to the expected, frequent, and repeated images. During this task, the activity of excitatory neurons as well as inhibitory SST and VIP neurons were recorded separately,

using 3 distinct mouse lines (see Methods). Neuronal cells within each line were simultaneously imaged at 4 cortical depths in 2 visual areas (V1 and higher visual area: LM; Fig. 1B,C; Supplementary Fig. 1).

We first studied how excitatory and inhibitory subtypes (VIP, SST) respond to image presentations and omissions, and how their responses depend on cortical area and depth. The population average of neural activities demonstrated a clear difference across cell classes (Fig. 1D,E). Excitatory neurons and, more robustly, SST neurons in all areas and layers were activated by expected images, i.e. the expected events. Conversely, excitatory neurons, overall, did not respond to omissions, except for those in the middle depth of LM. SST neuronal activity was slightly decreased after omissions in all recorded locations (Fig. 1D,E top, middle). In sharp contrast to excitatory and SST neurons, VIP neurons in all layers and areas were robustly activated after omissions (Fig. 1D,E bottom). Specifically, VIP cells demonstrated small ramping activity approximately 250 ms prior to each image presentation that further increased when the image was not presented (i.e. omission), and was immediately inhibited by image presentation, confirming previous results using the same behavioral task[9].

This multi-area dataset allowed comparison of neuronal responses across cortical depths of V1 and LM in each experiment (Fig. 1D,E). We found that excitatory neuronal responses to images were stronger in deeper layers of both visual areas, particularly V1. SST responses to images were strongest at middle depths of V1, corresponding to layer 2/3 and superficial layer 4, but did not differ among cortical layers of LM. VIP responses to neither images nor omissions were significantly different across cortical layers or areas, as reported previously[6], although the responses were slightly stronger in deeper layers. Therefore, in both V1 and LM, images are more robustly represented by excitatory and SST neurons, particularly in deeper layers; while unexpected omissions are represented by VIP neurons (Fig. 1D,E; Stats: Supplementary Table 1).

It is suggested that feedback pathways, from higher order cortical areas to superficial layers of V1, convey prediction signals[10,11,13]. In light of these studies, we used simultaneous multi-plane recording to investigate if cortical interactions might be different when expectations are violated, i.e. during omissions. As a proxy for cortical interactions, we measured pairwise noise correlations across repeated image presentations, as well as

omissions, for each cell type. Correlations were computed within each area (V1-V1, LM-LM; Supplementary Fig. 2-4), as well as across areas (V1-LM; Fig. 2). We also performed a shuffling analysis to control for firing rate impacts on correlations (Methods; Supplementary Fig. 2-4)

In the excitatory network, V1-V1 and V1-LM coactivations increased after image presentations, but only in deep layers (V1-LM: Fig. 2; V1-V1: Supplementary Fig. 2A,D). Interestingly, excitatory network coactivation in LM-LM did not change after images, across any pair of cortical layers (LM-LM: Supplementary Fig. 2B,E). Following omissions, there was no significant change in the coactivation of excitatory neurons across visual layers and areas (Fig. 2; Supplementary Fig. 2A,D).

In the SST network, images increased coactivation more broadly across visual layers and areas, particularly in our "middle" imaging depth, i.e., deep layer 2/3 and superficial layer 4 (Fig. 2; Supplementary Fig. 3). Omissions, on the other hand, slightly decreased coactivation of neurons relative to the baseline across most layers of the SST network (Fig. 2; Supplementary Fig. 3).

In the VIP network, coactivation patterns were strikingly different compared to other cell types: neurons in *superficial* layers of V1 and LM were strongly coactivated following omissions. There was also a weak coactivation prior to images in most layers and areas (Fig. 2; Supplementary Fig. 4). Notably, correlations were overall much weaker among excitatory neurons compared to inhibitory neurons (Fig. 2), confirming previous results and indicating stronger local connectivity between inhibitory neurons[14,15].

Our results suggest that the VIP network in the visual cortex is strongly recruited during unexpected events: cells from all layers of V1 and LM are active during omissions and neural correlations are especially strong in the superficial layers. One possible explanation is that prediction errors (i.e. mismatch between the animal's expectation of a given image and its omission) generate a shared input signal that co-activates VIP cells in superficial layers across areas, thereby synchronizing the local cortical network. An alternative hypothesis is that VIP activity during omissions may be due to co-occurring changes triggered by unexpected events. Such changes could include variations in animal's locomotion and arousal, as well as alterations in the representation

121 of environmental cues, or internal states. This hypothesis is corroborated by recent reports[16,17], including our
122 own[18], that have demonstrated a broad range of behavioral features modulate brain-wide network dynamics,
123 raising the possibility that our correlation findings could be contributed by global brain fluctuations driven by
124 widespread behavioral events. Importantly, in our previous work [9], we had found that the VIP population
125 carries no information about the identity of the previous image. These results demonstrated that VIP neurons
126 do *not* encode stimulus-specific prediction error signals; instead suggesting that VIP omission activity may
127 represent a broad non-specific surprise signal.

128 To address whether VIP correlated activity during omissions may be affected by global contextual changes,
129 we performed Tensor Component Analysis (TCA)[19] on real and trial-shuffled data, aligned on images and
130 omissions (number of sessions = 22, 15, 21 for excitatory, SST, and VIP cells, respectively). TCA allowed
131 decomposing neural population activity into distinct components, and trial shuffling allowed studying
132 contextual representation (see below). Each TCA component was characterized by weights indicating the
133 unique contributions from individual neurons, trials, and time points (Fig. 3A,B). Importantly, we defined TCA
134 components as image-specific or omission-specific if their trial weights were significantly higher on images
135 or omissions, respectively (ranksum test across image and omission weights; $p < 0.05$; Fig. 3E). Components
136 in which trial weights were not significantly different between images and omissions were excluded from
137 further analyses (number of non-specific components, pooled across sessions; mean +/- SD across 100 TCA
138 iterations; Excitatory: 38 +/- 4; SST: 20 +/- 3 ; VIP: 55 +/- 3).

139 To investigate the neural coding of images and omissions across visual areas and layers, we compared the
140 absolute magnitude of TCA neuronal weights among V1, LM, superficial and deep layers (imaging depth
141 smaller or bigger than 250μm, respectively; Fig 3C,D, red; for the details of TCA stats see Table S2). In
142 omission-specific components, VIP neurons had stronger weights in *superficial* layers of both V1 and LM.
143 This contrasted with excitatory and SST neurons which had stronger weights in deeper layers of V1 and LM
144 (Fig. 3C, red). In image-specific components, excitatory neurons in deeper layers of V1 and LM had stronger
145 weights; while SST neurons showed stronger weights in V1 compared to LM, independent of cortical layer

146 (Fig. 3D red). VIP neurons had few image-specific components, hence excluded from our analysis of neuronal
147 weights.

148 TCA neural weights for image- and omission-specific components reflect the coding strength of individual
149 neurons on image and omission trials, respectively. This neural coding could arise from the direct presentation
150 of images and omissions, or could be influenced by co-occurring behavioral and sensory features in adjacent
151 or remote areas, which we collectively term contextual features. To distinguish the neuronal coding of images
152 and omissions from these contextual influences, we performed TCA on trial shuffled data, for which trial
153 identities were shuffled *within* the image and omission conditions for each neuron (see Methods). This
154 shuffling procedure allowed us to preserve image- and omission-specific responses while removing any
155 contextual information that co-occurred across neurons during images and omissions. In contrast, in the non-
156 shuffle data, TCA components could encode the presence of images and omissions as well as any context-
157 dependent information present in the neuronal population.

158 Shuffling led to a general decrease of neuronal weight magnitudes for omission-specific components in all cell
159 types (Fig. 3C blue: shuffled; red: real data). This reduction was most prominent in VIP neurons in the
160 superficial layers, and absent from the VIP neurons in deep V1 (Fig. 3C, right). These findings suggest that
161 omissions are co-represented with contextual information in the visual cortex in excitatory, SST, and
162 particularly superficial VIP neurons. The major exception is that VIP neurons in the deeper layers of V1 show
163 a smaller contribution from contextual signals (Wicoxon ranksum test; $p=0.009$).

164 Interestingly, shuffling had a smaller effect on the neuronal weights of image-specific components in excitatory
165 and SST cells (Fig. 3D blue vs. red; Note the absence of VIP neurons from this analysis due to their few
166 numbers of image components). This finding suggests that during images, excitatory neurons primarily encode
167 image information, overriding contextual signals. This is aligned with previous studies that demonstrated
168 reduced neuronal variability during sensory inputs[20].

169 Indeed, all cell types, and particularly VIP neurons, had more TCA components that were omission-specific
170 vs. image-specific (Fig. 3E, black: images, gray: omissions). This finding suggests that omission responses in
171 our neuronal population exhibit more variability than image responses and is aligned with our TCA shuffled
172 analysis which demonstrated response variability during omissions was associated with the co-representation
173 of contextual features.

174 **DISCUSSION**

175 Recent work had demonstrated that the VIP population is functionally quite diverse, and strongly represents
176 behavioral features and unexpected image omissions[9]. Extending these findings, our current study
177 demonstrates that during unexpected omissions, VIP neurons in the visual cortex not only encode a non-
178 specific surprise signal but also a substantial amount of contextual information. This dual representation of
179 omissions and context is predominantly observed in superficial VIP neurons, with notably reduced context
180 coding in deeper layers.

181 Multiple lines of work have suggested the role of VIP neurons in encoding external and internal features. VIP
182 neurons can represent behavioral state, leading to gain modulation in sensory cortices[4,21]. VIP activity increases
183 with pupil dilation[22], locomotion[4,23], and whisking [3,24], leading to enhanced sensory modulation during high
184 attention and arousal states that follow surprising and reinforcement signals[4,21–23,25,26]. VIP neurons are driven
185 by neuromodulators, particularly cholinergic inputs[4], as well as long-range cortical and thalamic inputs[2,3]. VIP
186 neurons have also been shown to play a key role to surround suppression[27], when neuronal responses to
187 centered visual stimuli can be decreased by the presence of adjacent stimuli.

188 Our study suggests that the VIP network in primary and higher visual cortices comprises two subnetworks with
189 distinct roles, associated with superficial and deeper layers. In V1, the subnetwork in the deeper layers encodes
190 information about omissions, independently of context, while the subnetwork in the superficial layers
191 concurrently encodes both context and omission information. Importantly, the omission coding by VIP cells is
192 non-specific, lacking any information about which stimulus was omitted, as suggested by our recent work

193 (Garrett 2023). This supports the hypothesis that VIP neurons broadly encode surprise signals rather than
194 specific mismatches between expected and actual stimuli. Our findings suggest a potential specialization within
195 the VIP network that enhances the reliable representation of sensory features in deeper V1, while integrating
196 sensory and contextual features in the superficial layers of V1, hence facilitating adaptive responses to dynamic
197 environmental stimuli.

198 The ability of VIP cells to disinhibit excitatory neurons through its connectivity pattern, is ideally suited to
199 facilitate the propagation of selective activity patterns through the cortex, leading to attention towards
200 contextual sensory inputs[26] when needed, or the emergence of motor behaviors[28,29]. We propose a framework
201 in which VIP cells serve as gatekeepers of contextual information, which are dynamically activated when
202 transitioning from expected to unexpected stimuli (Fig. 4). This framework is in agreement with the numerous
203 studies on the activity of the inhibitory network. Future experiments should aim to simultaneously record from
204 all those cell types in the context of dynamic but controlled sensory uncertainty.

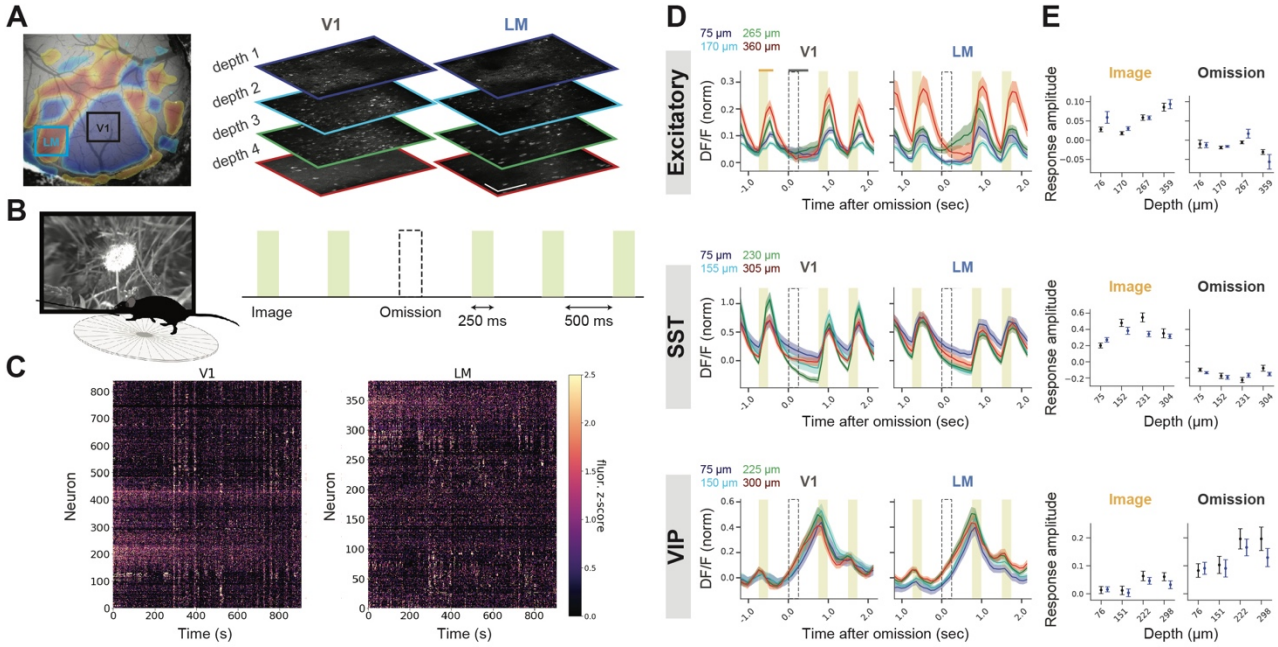


Figure 1. Excitatory and inhibitory cell types distinctly encode expected images and unexpected omissions across visual areas and layers. A) Allen Institute Visual Behavior dataset: calcium signals were simultaneously imaged across two cortical columns, including 4 depths (approximately 100 µm, 200 µm, 300 µm, and 375 µm), in visual areas V1 and LM, and in 3 distinct mouse lines tagged for excitatory, and 2 inhibitory cell types: VIP and SST. **B)** Mice, walking on a disk, view a familiar set of images, separated by a gray screen. Images are unexpectedly omitted on 5% of the presentation. **C)** Example population activity of excitatory neurons, recorded simultaneously from ~1180 neurons in V1 and LM. **D)** Population-averaged calcium responses to images (shaded rectangle) and omissions (dashed rectangle), across 2 visual areas (V1, LM), and 4 cortical depths (blue, cyan, green, red), in 3 different mouse lines (A-C: excitatory, SST, VIP). ΔF/F traces normalized to baseline standard deviation. Mean +/- SEM; n = 24, 22, 24 sessions for excitatory, SST, and VIP, respectively. **E)** Quantification of evoked responses averaged over 350ms after images, and 500ms after omissions (time-windows used for quantification of image and omission-evoked responses are indicated by orange and gray horizontal lines above top left panel).

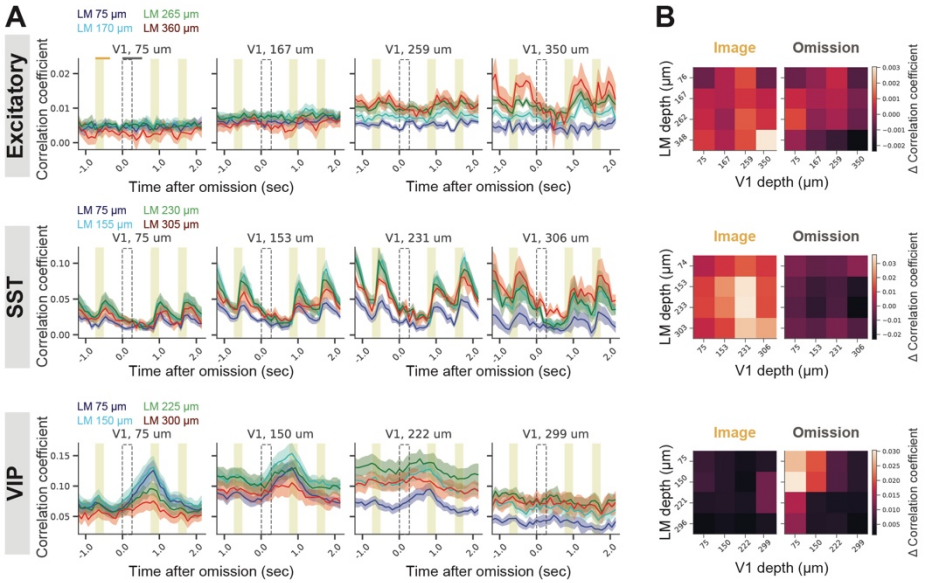


Figure 2. Distinct cortical interactions occur across primary and higher-order visual areas following expected and unexpected events. A) Spearman correlation coefficients between V1 and LM neurons at different cortical depths (colors indicate LM depths; each subplot corresponds to a given V1 depth). **B)** Change in correlation coefficients during images (left) and omissions (right) relative to baseline correlation coefficient. Correlation coefficients quantified over 500ms after images, and 750ms after omissions. Traces and heatmaps: mean +/- SEM; n = 8, 6, 9 mice for excitatory, SST, and VIP cell types, respectively.

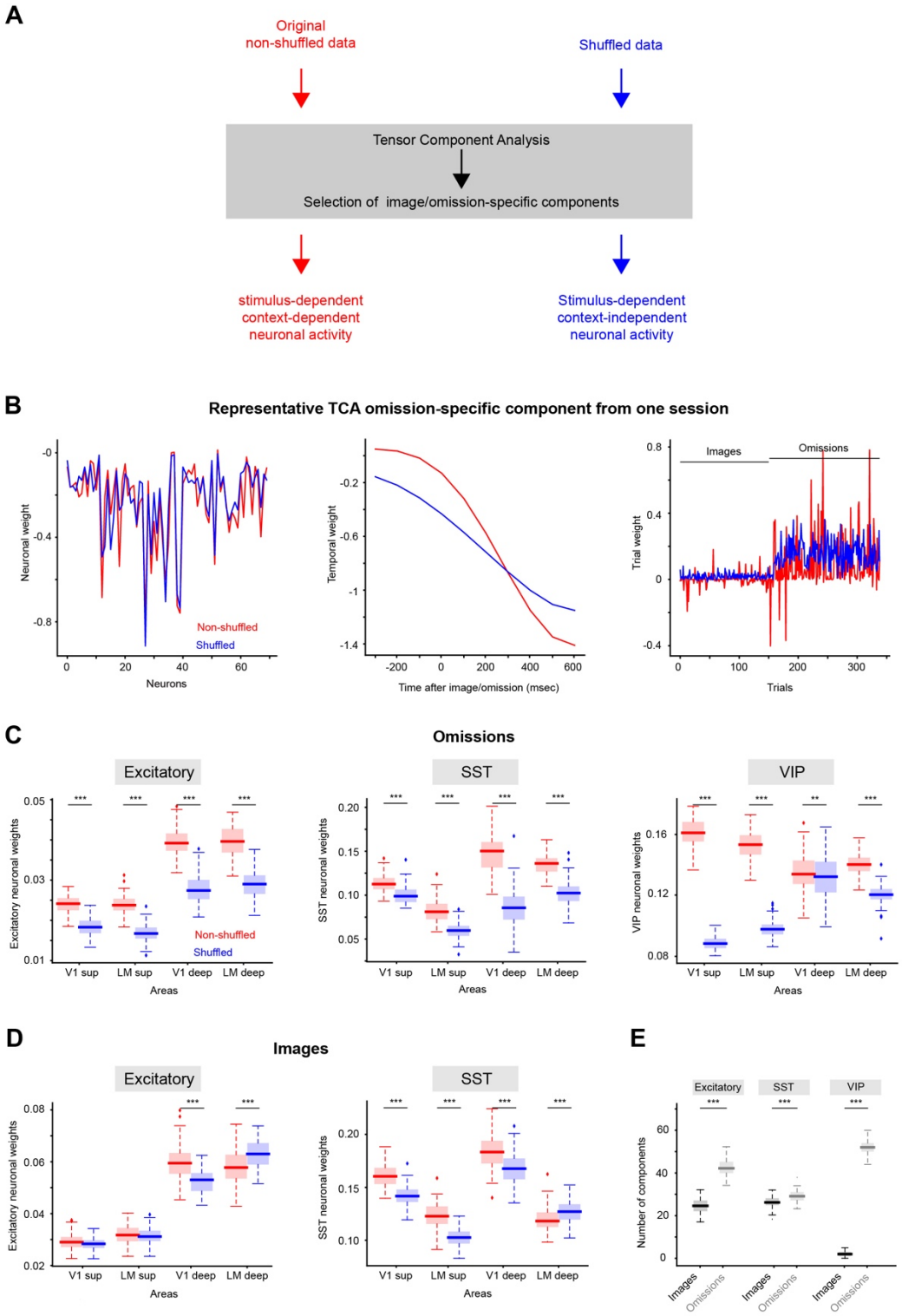


Figure 3. Unexpected events lead to enhanced context representation in excitatory and particularly VIP neurons in superficial layers. A) Schematic of the analysis pipeline. Imaging data were decomposed using Tensor Component Analysis (TCA; red: original; blue: shuffled). Shuffled data were obtained by permuting trials identities within each category (image and omission). Each TCA component was identified as image-

specific, omission-specific, or non-specific depending on the weight magnitude of the TCA trial component for images vs. omissions (Methods). **B)** Left to right: neuronal, temporal, and trials weights of a representative omission-specific TCA component (red: original; blue: shuffled). In this example, trials from 1 to 159 are images, while the remaining trials are omissions. TCA was performed on spikes extracted from DF/F traces aligned on images and omissions, including 300ms before to 700ms after images and omissions. **C)** Neuronal weight of the omission-specific components for excitatory, SST, and VIP neurons (left to right) extracted from original and shuffled data across areas (Wilcoxon ranksum test; all $p<0.001$ except for VIP V1 deep: $p=0.009$). **D)** Same as panel C for image-specific components for excitatory and SST neurons (left to right; Wilcoxon ranksum test; all $p<0.001$ except for excitatory V1 superficial: $p=0.8$ and excitatory LM superficial: $p=0.5$). **E)** Number of components pooled across sessions, for excitatory, SST, and VIP neurons (left to right) from original data compared between images and omissions (black and gray boxes, respectively; Wilcoxon ranksum test; all $p<0.001$).

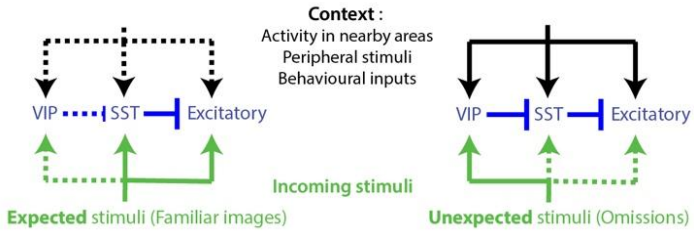


Figure 4. Proposed mechanism: Triggered by unexpected events, VIP neurons gate contextual signal flow across cortical cell types and areas. Left) Expected stimuli, such as familiar images in our study, are encoded by excitatory and SST neurons. **Right)** When unexpected stimuli occur, a non-specific surprise signal activates VIP cells to promote contextual integration across all cell types. This contextual information may originate from nearby and distant brain areas, encoding a variety of signals, including internal (cognitive and motivational states), or external (sensory and behavioral) information. The presence of this contextual information facilitates the updating of the unexpected stimulus representation in the local network.

256 **ONLINE METHODS**

257 **Processing of calcium imaging movies**

258 The preprocessing of all calcium imaging data was done within the imaging processing pipeline described in
259 detail elsewhere[34] and crosstalk removal was applied after pre-processing. Briefly, all data was first corrected
260 for brain motion by performing rigid registration in two dimensions. Then, cell-segmentation was performed
261 to identify spatial masks of active neurons. Further, fluorescence from spatially-overlapping neuronal masks
262 was unmixed and corrected for neuropil contamination. Finally, mask-matching and crosstalk removal was
263 performed using FastICA[37] and ghost cells were filtered, and lastly, $\Delta F/F$ was computed on corrected masks.

264 **Description of datasets**

265 We used the publicly available Allen Institute Visual behavior dataset: https://portal.brain-map.org/circuits-
266 behavior/visual-behavior-2p. Details are explained in our previous work (Garrett 2023). In brief, this dataset
267 included 3 mouse lines: excitatory mouse line: Slc17a7-IRES2-Cre ; Camk2a-tTA ; Ai93(TITL-GCaMP6f),
268 inhibitory subpopulation VIP: Vip-IRES-Cre ; Ai148 (TIT2L-GC6f-ICL-tTA2), and inhibitory subpopulation
269 SST: Sst-IRES-Cre ; Ai148 (TIT2L-GC6f-ICL-tTA2). For the excitatory cell line, 24 sessions from 8 mice;
270 for the SST cell line, 22 sessions from 6 mice; for the VIP cell line, 24 sessions from 9 mice were recorded.

271 **Population averages of neuronal responses**

272 Traces were aligned to the onset of image omission; since images were presented regularly, omission alignment
273 also aligned the traces on images. The median of neural responses was computed across trials. The average
274 response was then computed across neurons, for each session. A grand average was then computed across
275 sessions (see Fig. 1D, traces). To quantify the image-evoked responses, the calcium trace of each neuron was
276 averaged over 350ms after image onset for excitatory and SST neurons, and over [-250, 100]ms relative to
277 image onset for VIP neurons to account for their anticipatory response. Response quantification was done on
278 the mean (across trials) trace of each neuron. The same quantification was performed for omission-evoked

279 responses, except a 500ms window was used for quantification. The responses were averaged across neurons,
280 for each session, and then a grand average was computed across sessions (see Fig. 1E; error-bars).

281 **Correlation of neural responses across cortical planes**

282 On omission-aligned traces, Spearman correlation coefficient was computed between pairs of neurons across
283 trials. Here, by trial we mean omissions. Correlations were computed for every individual frame, over [-1,
284 2] sec relative to the omission. This procedure was done for all pairs of neurons; then an average value was
285 computed across all pairs (Fig. 2). Neuron pairs were present within the same plane, or in 2 different planes.
286 This analysis allowed studying how the response of neurons (within the same plane or in different areas/layers)
287 covaried across trials, and how this coactivation changed at different moments (e.g. after images vs. omissions).
288 Spearman correlation coefficients were also computed on shuffled traces (Supplementary Fig. 2-4 d-f), which
289 were driven by independently shuffling trial orders for each neuron. For each neuron pair, shuffling was
290 repeated 50 times, resulting in a distribution of correlation coefficients for shuffled data. This shuffling analysis
291 controlled for firing rate impacts on our correlation results.

292 To compute "noise" correlations, we measured pairwise correlations of "signal"-removed traces. "Signal" was
293 computed by taking the average neural response to each image type (there were 8 images in each session), and
294 subtracting the average from the response of individual trials of that image type.

295 To quantify coactivation of neurons (Fig. 2, heatmaps; Supplementary Fig. 2-4d-f), correlations coefficients
296 ("cc") were first averaged over 500 ms after images for excitatory and SST neurons, and over [-250, 250] ms
297 relative to image onset for VIP neurons, accounting for their anticipatory response. Then, correlation
298 coefficients were averaged across baseline frames. We call this quantity "image cc". To quantify omission-
299 evoked coactivation, we averaged correlation coefficients over 750 ms after omissions. We call this quantity
300 "omission cc". Next, we quantified baseline coactivation by averaging correlation coefficients across baselines
301 frames. Baseline was defined as the frame immediately preceding each image presentation, for excitatory and
302 SST neurons, and 250 ms earlier than each image presentation for VIP neurons. We call baseline quantification

303 of correlation "baseline cc". Finally, we measured the change in coactivation during images or omissions by
304 subtracting out "baseline cc" from "image cc", or "omission cc". This quantity is plotted in Fig. 2, heatmaps,
305 and Supplementary Fig. 2-4d-f, error-bars.

306 The Python package scipy (scipy.stats.spearmanr) was used for computing correlations, and p-values were
307 computed in 2 ways: 1) using the p-value output of the spearmanr package; 2) manually computing the p-value
308 by comparing the correlation coefficient of real (non-shuffled) data with the shuffled distribution using 2-sided,
309 1-sample t-test (using the Python package scipy.stats.ttest_1samp).

310 **Tensor Component Analysis**

311 Tensor Component Analysis (TCA) was performed on inferred spikes from the DF/F trace using the tensortool
312 package (https://github.com/neurostatslab/tensortools). TCA was set to extract 5 components from the trials
313 split from -300 to +600 msec relative to the image or omission onset. The resulting components were first
314 filtered according to their image- or omission-specificity. Specifically, we compared the weights of image and
315 omission trials for each component (Wilcoxon ranksum test, alpha <0.05) and identified as condition-specific
316 the components showing a significant difference between the two. On this subset of components, those with
317 larger absolute average weights during image trials were classified as image-specific, while those with larger
318 absolute average weights during omission trials were classified as omission-specific. To assess the variability
319 across trials, on each subset of components we computed the standard deviation across trial weights. This entire
320 procedure was performed 100 times to account for the variability between TCA decompositions. To obtain a
321 null distribution that discarded context information while preserving image- and omission-specific
322 information, we shuffled trials for each neuron keeping separated image and omission trials (e.g. the activity
323 of an omission trial was reassigned to another omission trial independently for each neuron). Similarly to the
324 original data, the procedure was performed 100 times to account for variability in shuffling and TCA
325 decomposition. While TCA explained a relatively small amount of the overall signal, especially in the case of
326 excitatory neurons, the components were highly reproducible across iterations (for TCA error and similarity

across iterations see Table S3). As an additional control, we performed the same analysis with a larger number of reconstructed components by TCA and found similar results (Supplementary Fig. 5).

Statistical tests

We used two-way ANOVA, followed by Tukey HSD to compare population averages across cortical layers. Two-sided t-test was used to compare correlations between real and shuffled data, for each cortical plane. A p-value of 0.05 was used as the significance threshold. For comparison of the correlation coefficients, we used two-tailed, two-sample Kolmogorov–Smirnov tests.

Data availability

Data are publicly available as Allen Institute Visual Behavior Dataset: https://portal.brain-map.org/circuits-behavior/visual-behavior-2p

Code availability

Crosstalk removal was performed using custom routines employing FastICA which is available as part of scikit-learn python package (https://scikit-learn.org/). Code to perform population average and correlation analyses are available at: https://github.com/farznaj/multiscope_fn. Code to perform TCA and the related statistical tests are available at: https://github.com/SimoneRusso/Multiscope_TCA

Acknowledgements

We wish to thank the founder of the Allen Institute for Brain Science, Paul G. Allen, for his vision, encouragement and support.

348 **Competing interests**

349 The dual-beam add-on module (J.L.) intellectual property has been licensed to Thorlabs. Inc., by the Allen
350 Institute. S.R. is the Chief Medical Officer of Manava Plus.

351 **Correspondence and requests** for materials should be addressed to J.L, and F.N..

352

353 **References**

354 1. Gouwens, N. W. *et al.* Integrated Morphoelectric and Transcriptomic Classification of Cortical
355 GABAergic Cells. *Cell* **183**, 935–953.e19 (2020).
356 2. Wall, N. R. *et al.* Brain-Wide Maps of Synaptic Input to Cortical Interneurons. *J. Neurosci.* **36**, 4000–
357 4009 (2016).
358 3. Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G. & Rudy, B. A disinhibitory circuit mediates motor
359 integration in the somatosensory cortex. *Nat. Neurosci.* **16**, 1662–1670 (2013).
360 4. Fu, Y. *et al.* A cortical circuit for gain control by behavioral state. *Cell* **156**, 1139–1152 (2014).
361 5. Bastos, G. *et al.* Top-down input modulates visual context processing through an interneuron-specific
362 circuit. *Cell Rep.* **42**, 113133 (2023).
363 6. Garrett, M. *et al.* Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. *Elife*
364 **9**, (2020).
365 7. Hertäg, L. & Sprekeler, H. Learning prediction error neurons in a canonical interneuron circuit. *Elife* **9**,
366 (2020).
367 8. Hamm, J. P. & Yuste, R. Somatostatin Interneurons Control a Key Component of Mismatch Negativity
368 in Mouse Visual Cortex. *Cell Rep.* **16**, 597–604 (2016).
369 9. Garrett, M. *et al.* Stimulus novelty uncovers coding diversity in visual cortical circuits. *bioRxiv*
370 2023.02.14.528085 (2023) doi:10.1101/2023.02.14.528085.
371 10. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron*
372 **100**, 424–435 (2018).
373 11. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
374 12. de Vries, S. E. J. *et al.* A large-scale standardized physiological survey reveals functional organization
375 of the mouse visual cortex. *Nat. Neurosci.* **23**, 138–151 (2020).
376 13. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some
377 extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
378 14. Najafi, F. *et al.* Excitatory and Inhibitory Subnetworks Are Equally Selective during Decision-Making

379 and Emerge Simultaneously during Learning. *Neuron* **105**, 165–179.e8 (2020).
380 15. Packer, A. M. & Yuste, R. Dense, unspecific connectivity of neocortical parvalbumin-positive
381 interneurons: a canonical microcircuit for inhibition? *J. Neurosci.* **31**, 13260–13271 (2011).
382 16. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and
383 engagement across the mouse brain. *Nature* **576**, 266–273 (2019).
384 17. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics
385 are dominated by richly varied movements. *Nat. Neurosci.* **22**, 1677–1686 (2019).
386 18. Ramadan, M. *et al.* A Standardized Nonvisual Behavioral Event Is Broadcasted Homogeneously across
387 Cortical Visual Areas without Modulating Visual Responses. *eNeuro* **9**, (2022).
388 19. Williams, A. H. *et al.* Unsupervised Discovery of Demixed, Low-Dimensional Neural Dynamics across
389 Multiple Timescales through Tensor Component Analysis. *Neuron* **98**, 1099–1115.e8 (2018).
390 20. Churchland, M. M. *et al.* Stimulus onset quenches neural variability: a widespread cortical phenomenon.
391 *Nat. Neurosci.* **13**, 369–378 (2010).
392 21. Pi, H.-J. *et al.* Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521–524 (2013).
393 22. Reimer, J. *et al.* Pupil fluctuations track fast switching of cortical states during quiet wakefulness.
394 *Neuron* **84**, 355–362 (2014).
395 23. Pakan, J. M. *et al.* Behavioral-state modulation of inhibition is context-dependent and cell type specific
396 in mouse visual cortex. *Elife* **5**, (2016).
397 24. Yu, J., Hu, H., Agmon, A. & Svoboda, K. Recruitment of GABAergic Interneurons in the Barrel Cortex
398 during Active Tactile Behavior. *Neuron* **104**, 412–427.e4 (2019).
399 25. Letzkus, J. J. *et al.* A disinhibitory microcircuit for associative fear learning in the auditory cortex.
400 *Nature* **480**, 331–335 (2011).
401 26. Karnani, M. M. *et al.* Opening Holes in the Blanket of Inhibition: Localized Lateral Disinhibition by
402 VIP Interneurons. *J. Neurosci.* **36**, 3471–3480 (2016).
403 27. Keller, A. J. *et al.* A Disinhibitory Circuit for Contextual Modulation in Primary Visual Cortex. *Neuron*
404 **108**, 1181–1193.e8 (2020).

28. Arroyo, S., Barati, S., Kim, K., Aparicio, F. & Ganguly, K. Emergence of preparatory dynamics in VIP interneurons during motor learning. *Cell Rep.* **42**, 112834 (2023).

29. Ren, C. *et al.* Global and subtype-specific modulation of cortical inhibitory neurons regulated by acetylcholine during motor learning. *Neuron* **110**, 2334–2350.e8 (2022).

Supplementary material

Excitatory, V1 image responses	one-way ANOVA: F= 16.3, p<0.001; TUKEY HSD: p<0.01 for all pairwise comparisons except layer 1-2/3 and layer 1-4.
Excitatory, LM image responses	TUKEY HSD: p>0.05 for all pairwise comparisons, except for layer 2/3-5: p=0.001.
SST, V1 image responses	one-way ANOVA: F=12.5, p<0.001; TUKEY HSD: p<0.05 for layer1-2/3, layer 1-4, and layer 4-5 pairwise comparisons.
SST, LM image responses	one-way ANOVA: F= 1.9, p=0.14.
VIP, V1 image responses	one-way ANOVA: F=3.0, p=0.03; yet TUKEY HSD: p>0.05 for all pairwise comparisons.
VIP, LM image responses	one-way ANOVA: F=2.4, p=0.07. VIP, V1 omission responses: one-way ANOVA: F=3.0, p=0.03; yet TUKEY HSD: p>0.05 for all pairwise comparisons. VIP, LM omission responses: one-way ANOVA: F=1.6, p=0.19.

Supplementary Table 1. Statistical tests corresponding to Fig. 1.

SLC - images		
ANOVA		
statistic=974.9242425360012, pvalue=2.053417877366087e-182		
Posthoc		
Group1	Group2	p
0	1	<0.001
0	2	0.0045
0	3	<0.001
1	2	<0.001
1	3	0.17
2	3	<0.001
Ranksum		
Group	p	effect size
1	0.5	-0.000559821
2	<0.001	0.005208667
3	0.8	-0.000677232
4	<0.001	-0.006407717

SLC - omissions		
ANOVA		
statistic=924.6397775278477, pvalue=2.0365188317079515e-178		
Posthoc		
Group1	Group2	p
0	1	<0.001
0	2	0.9
0	3	<0.001
1	2	<0.001
1	3	0.9
2	3	<0.001
Ranksum		
Group	p	effect size
1	<0.001	-0.007043246
2	<0.001	-0.010607699
3	<0.001	-0.005858144
4	<0.001	-0.011777935

SST - images		
ANOVA		
statistic=562.5678946813928, pvalue=2.34542498531712e-142		
Posthoc		
Group1		
0	1	0.26
0	2	<0.001
0	3	<0.001
1	2	<0.001
1	3	<0.001
2	3	<0.001
Ranksum		
Group	p	effect size
1	<0.001	-0.020148597
2	<0.001	0.008925503
3	<0.001	-0.018708875
4	<0.001	-0.015717922

SST - omissions		
ANOVA		
statistic=388.69757476891175, pvalue=1.3433542133282506e-117		
Posthoc		
Group1		
0	1	<0.001
0	2	<0.001
0	3	<0.001
1	2	<0.001
1	3	<0.001
2	3	<0.001
Ranksum		
Group		
1	<0.001	-0.021273879
2	<0.001	-0.032506687
3	<0.001	-0.010707941
4	<0.001	-0.061329306

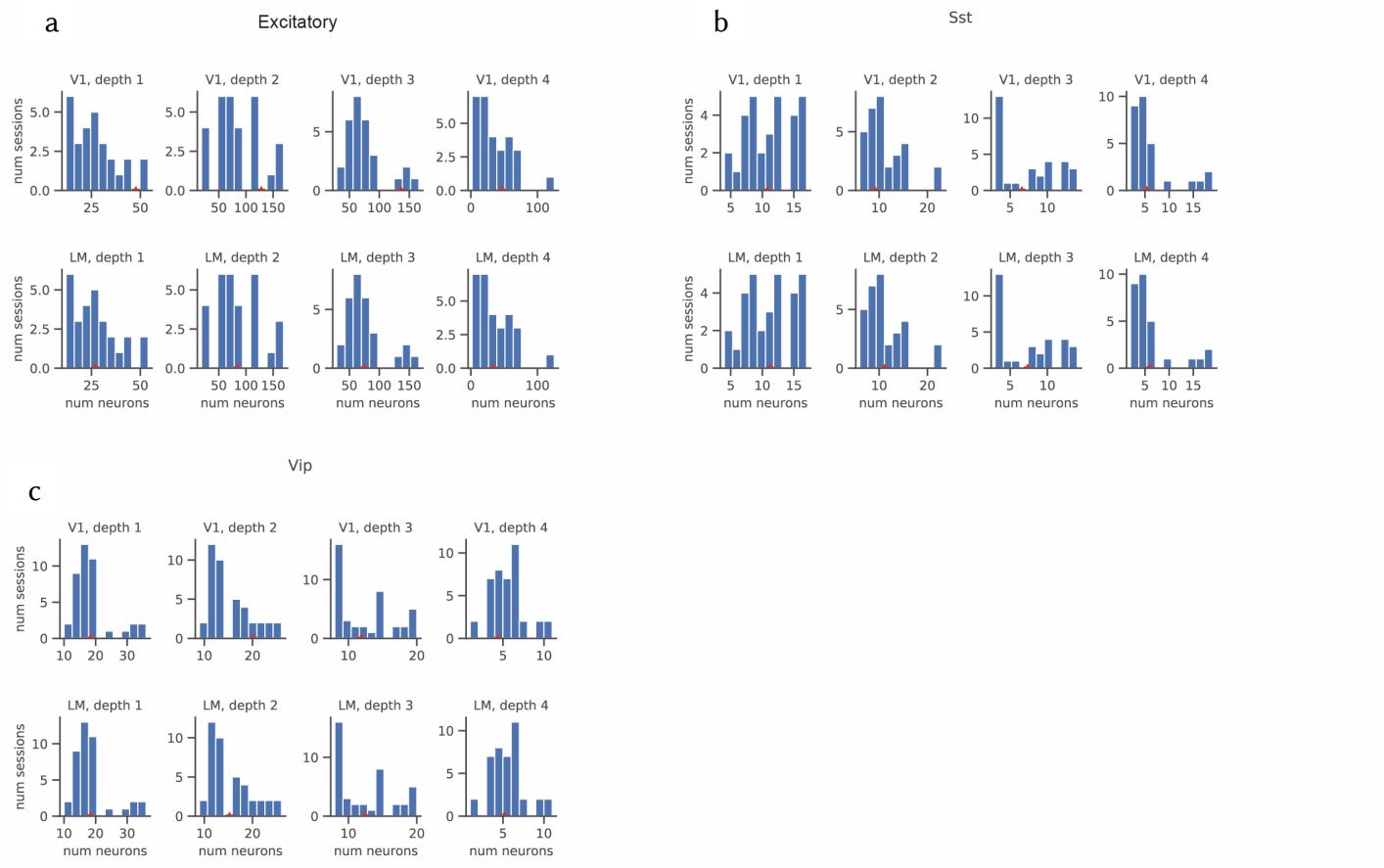
LEGEND	
group 1	LM superficial
group 2	LM deep
group 3	V1 superficial
group 4	V1 deep

VIP - omissions		
ANOVA		
statistic=152.09786442559385, pvalue=1.4251507357988139e-65		
Posthoc		
Group1		
0	1	<0.001
0	2	<0.001
0	3	<0.001
1	2	<0.001
1	3	0.0058
2	3	<0.001
Ranksum		
Group		
1	<0.001	-0.054864618
2	<0.001	-0.017996513
3	<0.001	-0.068608776
4	0.009	-0.005409973

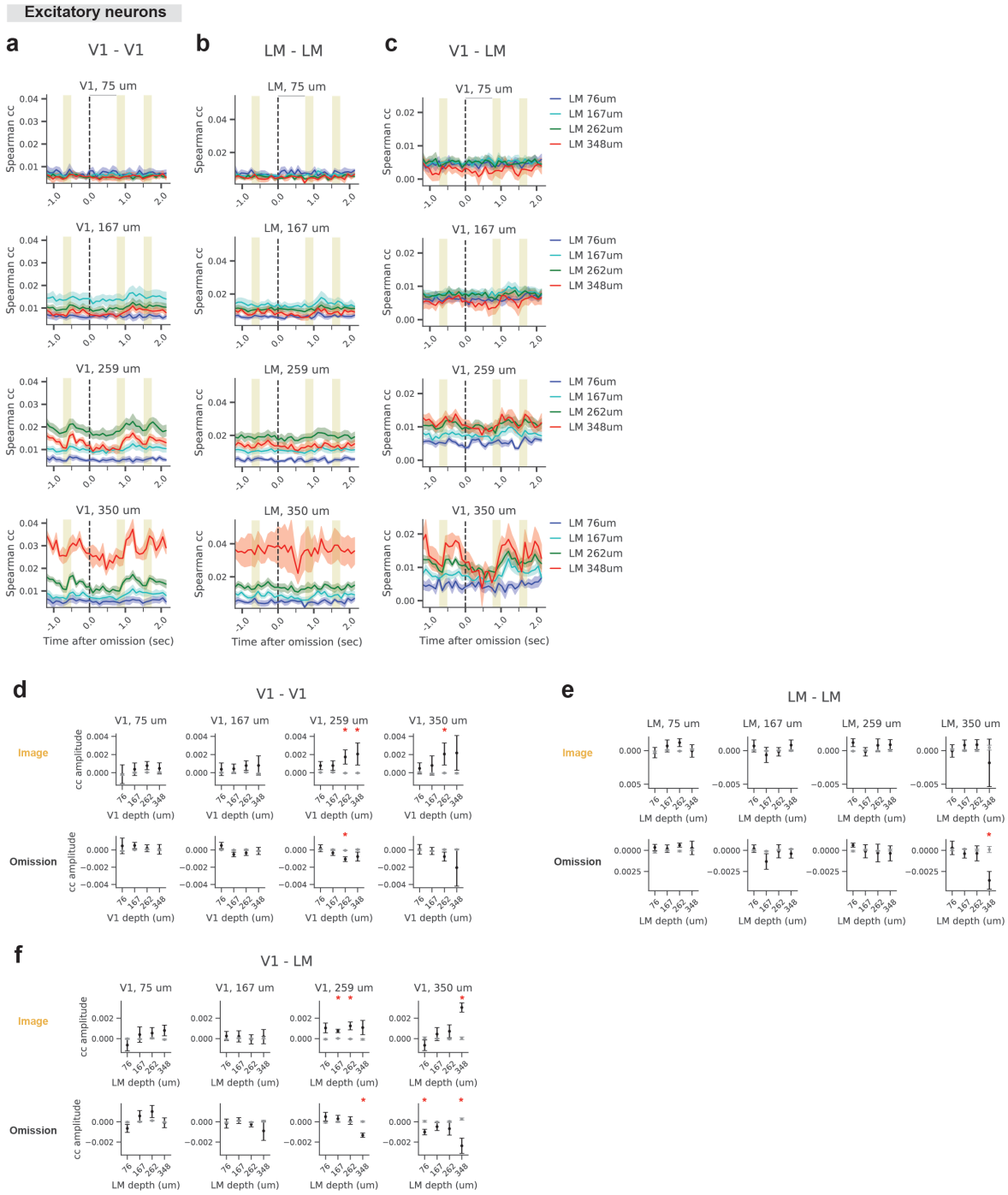
Supplementary Table 2. Statistical tables for TCA analysis.

Number of components	Celltype	Data	Error	Similarity
5	Excitatory	Original	0.9401	0.8879
		Shuffled	0.9681	0.8673
	SST	Original	0.7695	0.9227
		Shuffled	0.8365	0.9067
	VIP	Original	0.774	0.918
		Shuffled	0.8838	0.8886
10	Excitatory	Original	0.9084	0.8265
		Shuffled	0.9366	0.8038
	SST	Original	0.67	0.8695
		Shuffled	0.7481	0.858
	VIP	Original	0.6991	0.8515
		Shuffled	0.8238	0.8195

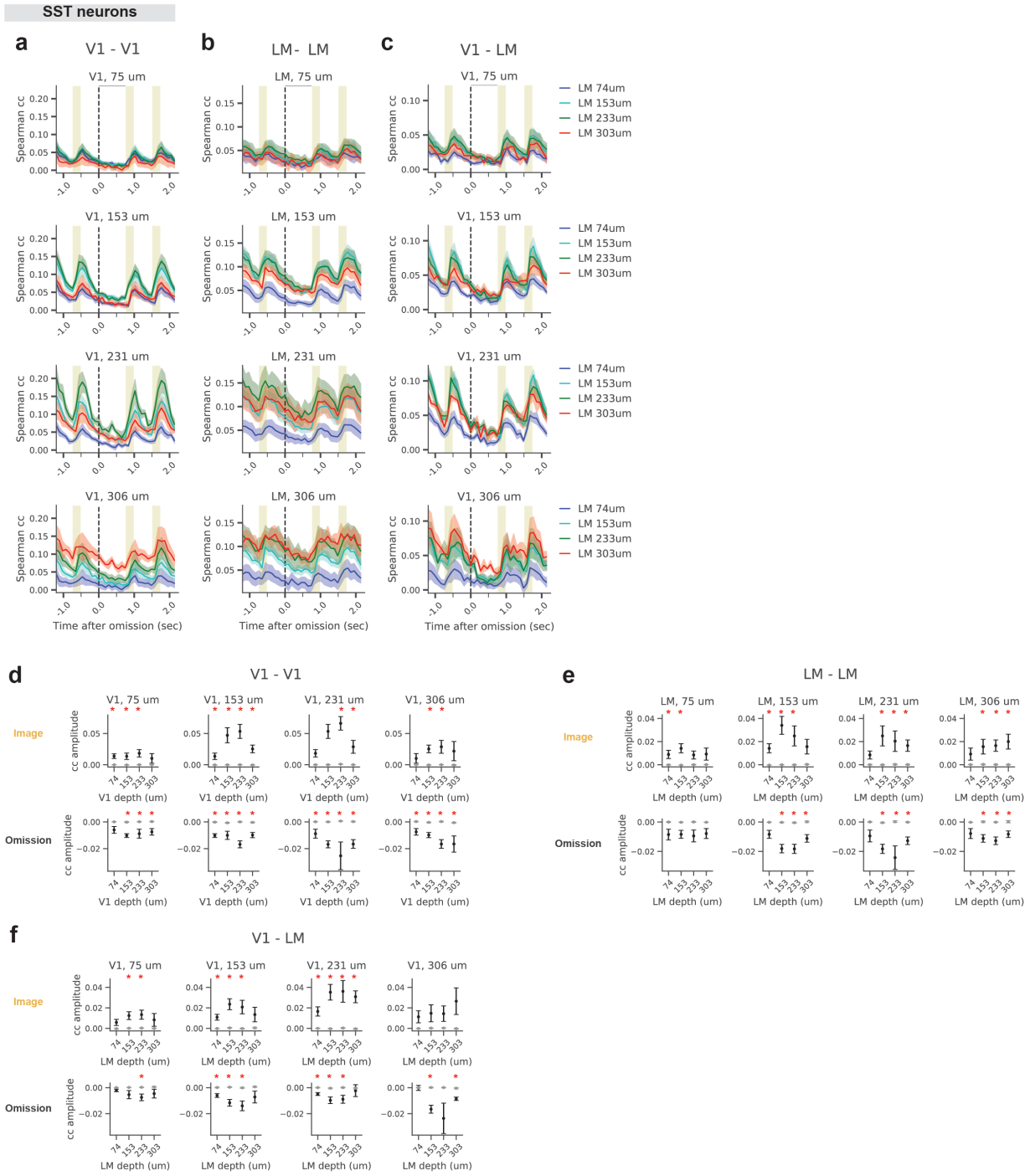
Supplementary Table 3. TCA error and similarity values across with across number of components, celltypes, and original vs shuffled data.



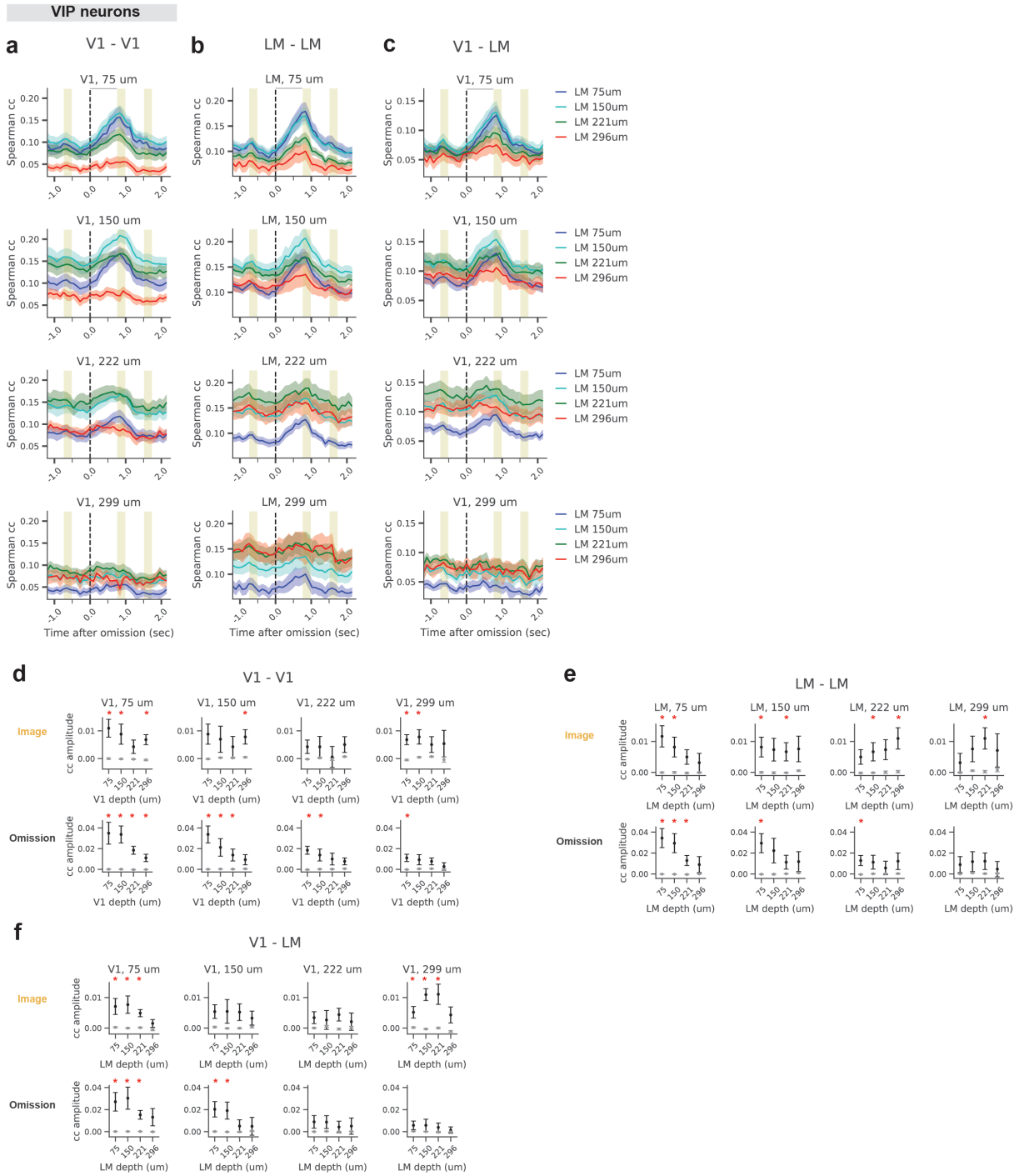
Supplementary Figure 1. Number of neurons per depth and area, for each cell type. Distribution of number of neurons recorded simultaneously across depths of V1 (top) and LM (bottom), in all recording sessions, for excitatory **(a)**, SST **(b)**, and VIP **(c)** mice. Note that our goal in this study was to chronically record the activity of distinct excitatory and inhibitory cell types, from the same field of view and across multiple sessions. Like in other studies involving the original Mesoscope[1,2], our system can record from a much larger number of neurons, however, our biological experiments were not designed for that purpose.



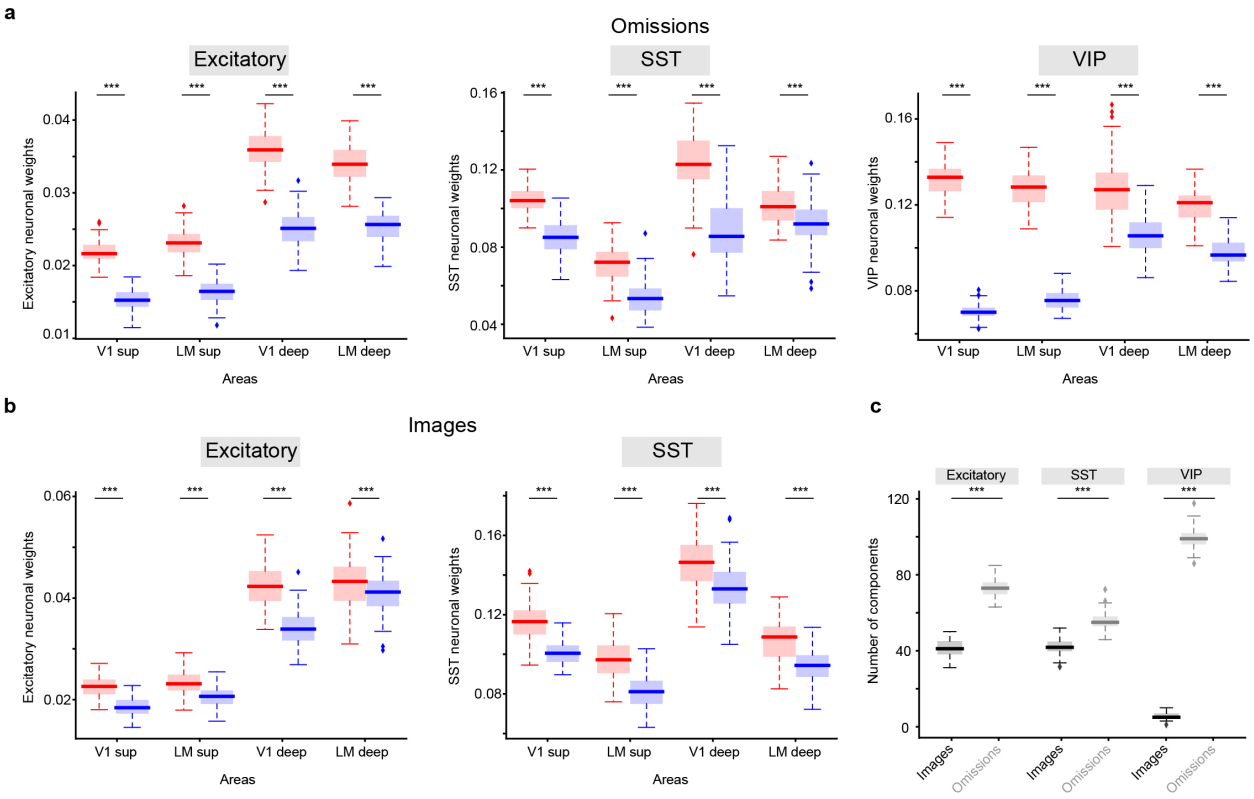
Supplementary Figure 2. Correlation of excitatory neurons within V1, within LM, and across V1-LM, during images and omissions. a-c Spearman correlation coefficients computed, at different moments in the trial, between activities of excitatory neurons located in different depths of V1-V1 (a), LM-LM (b), and V1-LM (c). **d-f** Change in correlation coefficients during images (top) and omissions (bottom) relative to the baseline correlation coefficient, computed for the real data (black), and trial shuffled data (gray). Correlation coefficients were quantified over 500 ms after images, and 750 ms after omissions. Red stars indicate statistical significance (two-sided t-test, real vs. shuffle data; $p<0.05$). Traces and error-bars: mean +/- SEM; n = 8 mice.



Supplementary Figure 3. Correlation **of SST neurons within V1, within LM, and across V1-LM, during images and omissions. a-c** Spearman correlation coefficients computed, at different moments in the trial, between activities of excitatory neurons located in different depths of V1-V1 (a), LM-LM (b), and V1-LM (c). **d-f** Change in correlation coefficients during images (top) and omissions (bottom) relative to the baseline correlation coefficient, computed for the real data (black), and trial shuffled data (gray). Correlation coefficients were quantified over 500 ms after images, and 750 ms after omissions. Red stars indicate statistical significance (two-sided t-test, real vs. shuffle data; $p<0.05$). Traces and error-bars: mean +/- SEM; n = 6 mice.



Supplementary Figure 4. Correlation of VIP neurons within V1, within LM, and across V1-LM, during images and omissions. a-c Spearman correlation coefficients computed, at different moments in the trial, between activities of excitatory neurons located in different depths of V1-V1 (a), LM-LM (b), and V1-LM (c). **d-f** Change in correlation coefficients during images (top) and omissions (bottom) relative to the baseline correlation coefficient, computed for the real data (black), and trial shuffled data (gray). Correlation coefficients were quantified over 500 ms after images, and 750 ms after omissions. Red stars indicate statistical significance (two-sided t-test, real vs. shuffle data; $p<0.05$). Traces and error-bars: mean +/- SEM; n = 9 mice.



Supplementary Figure 5. TCA with 10 components. a, b, and c depict the results for the TCA analysis for 10 components corresponding to Figure 3D,E,F, respectively.