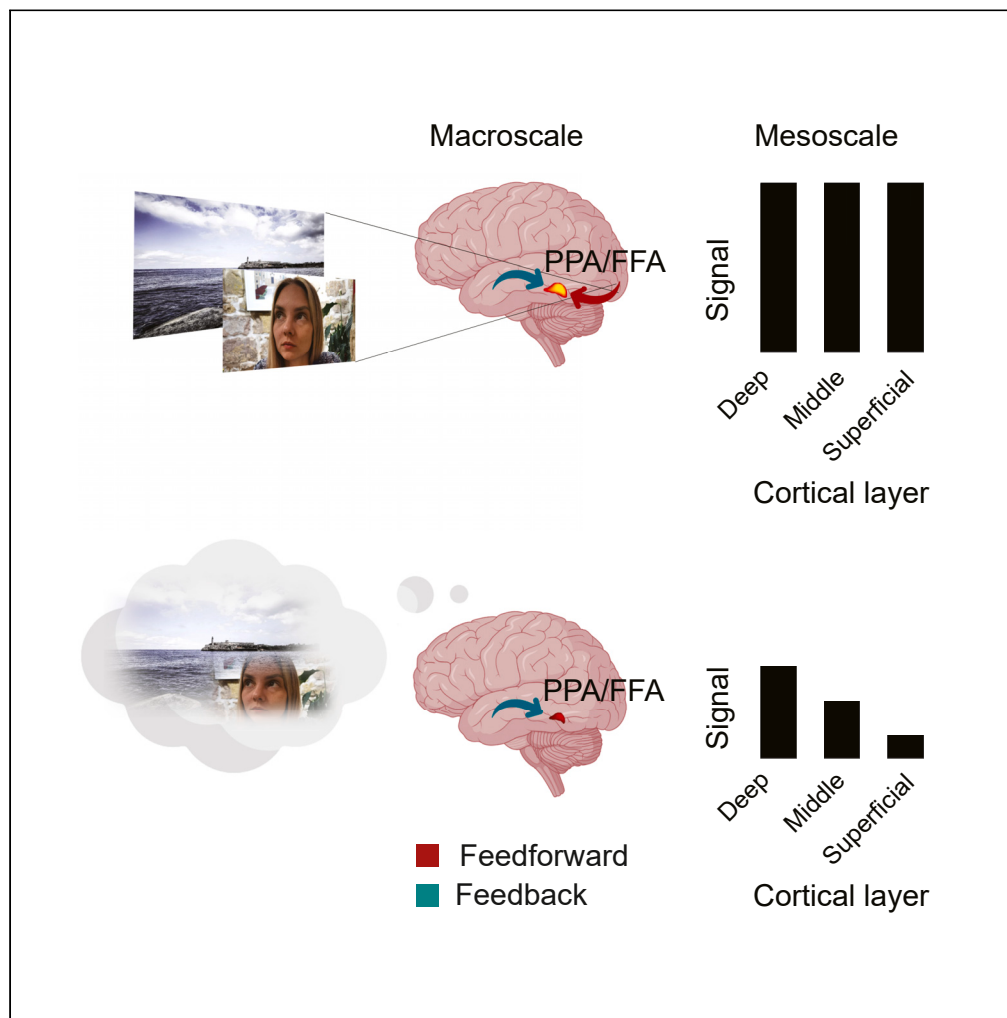


Article

Laminar dissociation of feedforward and feedback in high-level ventral visual cortex during imagery and perception



Tony Carricarte,
Polina
Iamshchinina,
Robert Trampel,
Denis Chaimow,
Nikolaus
Weiskopf,
Radoslaw M. Cichy

tcarricarte@gmail.com

Highlights

Visual imagery and perception have distinctive laminar responses in FFA and PPA

Carricarte et al., iScience 27, 110229
July 19, 2024 © 2024 The Author(s). Published by Elsevier Inc.
<https://doi.org/10.1016/j.isci.2024.110229>



Article

Laminar dissociation of feedforward and feedback in high-level ventral visual cortex during imagery and perception

Tony Carricarte,^{1,2,3,8,*} Polina Iamshchinina,⁴ Robert Trampel,⁵ Denis Chaimow,⁵ Nikolaus Weiskopf,^{5,6} and Radoslaw M. Cichy^{1,2,3,7}

SUMMARY

Visual imagery and perception share neural machinery but rely on different information flow. While perception is driven by the integration of sensory feedforward and internally generated feedback information, imagery relies on feedback only. This suggests that although imagery and perception may activate overlapping brain regions, they do so in informationally distinctive ways. Using lamina-resolved MRI at 7 T, we measured the neural activity during imagery and perception of faces and scenes in high-level ventral visual cortex at the mesoscale of laminar organization that distinguishes feedforward from feedback signals. We found distinctive laminar profiles for imagery and perception of scenes and faces in the parahippocampal place area and the fusiform face area, respectively. Our findings provide insight into the neural basis of the phenomenology of visual imagery versus perception and shed new light into the mesoscale organization of feedforward and feedback information flow in high-level ventral visual cortex.

INTRODUCTION

Visual imagery and perception are phenomenologically related, suggesting that they share neural machinery. Previous research confirmed this assumption at the macroscale of human cortical organization in category-selective regions in ventral visual cortex.¹ However, imagery and perception differ fundamentally in the information flow that underlies them²: perception is driven by the integration of sensory feedforward and internally generated feedback information,^{3,4} whereas imagery depends on feedback only.⁵ This suggests that, while visual imagery and perception may activate common cortical regions, they do so in fundamentally distinctive ways.

While feedforward has been consistently associated with middle layers,^{6–8} the interaction of feedback with cortical layers has exhibited greater variability, with feedback arriving in deep layers,^{9,10} superficial layers,^{7,8,11,12} or both of them.^{13,14} Notably, recent works suggest that this variability might be linked to two different modes of feedback^{9,15}: externally and internally generated feedback. Externally generated feedback originates simultaneously with the presentation of a stimulus, such as during attention directed toward a target,⁷ contextual filling-in,¹² and visual illusion,⁹ or in response to task demands.¹¹ In contrast, internally generated feedback primarily originates from internal cognitive processes that do not require external stimulation at the moment of initiation. This includes processes like memory recall,¹⁶ mental imagery,⁹ and expectation.¹⁰ While externally generated feedback has been primarily linked to the superficial layers, internally generated feedback has been mainly associated with the deep layers; however see Kok et al.¹⁷

In line with this, we hypothesized that feedforward and feedback information flow in high-level ventral visual regions differs at the mesoscale level of laminar organization (Figure 1A). Specifically, we predicted a uniform activation across cortical layers in perception, following empirical observations in V1.^{9,10,12} In contrast, in imagery we predicted a selective activation of the deep layers.

To investigate, we resolved neural activity during imagery and perception in the ventral visual cortex at the level of laminar organization that anatomically^{18,19} and functionally^{20,21} distinguishes feedforward from feedback information flow. We found distinctive laminar profiles for imagery and perception of scenes and faces in parahippocampal place area (PPA) and fusiform face area (FFA), respectively. Our findings provide insight into the neural basis of the phenomenology of visual imagery versus perception and shed new light on how feedforward and feedback information processing in high-level ventral visual cortex orchestrates human object vision.

¹Department of Education and Psychology, Freie Universität Berlin, 14195 Berlin, Germany

²Einstein Center for Neurosciences Berlin, Charité - Universitätsmedizin Berlin, 10117 Berlin, Germany

³Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität zu Berlin, 10117 Berlin, Germany

⁴Princeton Neuroscience Institute, Princeton University, New Jersey 08544, USA

⁵Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

⁶Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth Sciences, Universität Leipzig, 04103 Leipzig, Germany

⁷Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, 10117 Berlin, Germany

⁸Lead contact

*Correspondence: tcarricarte@gmail.com

<https://doi.org/10.1016/j.isci.2024.110229>



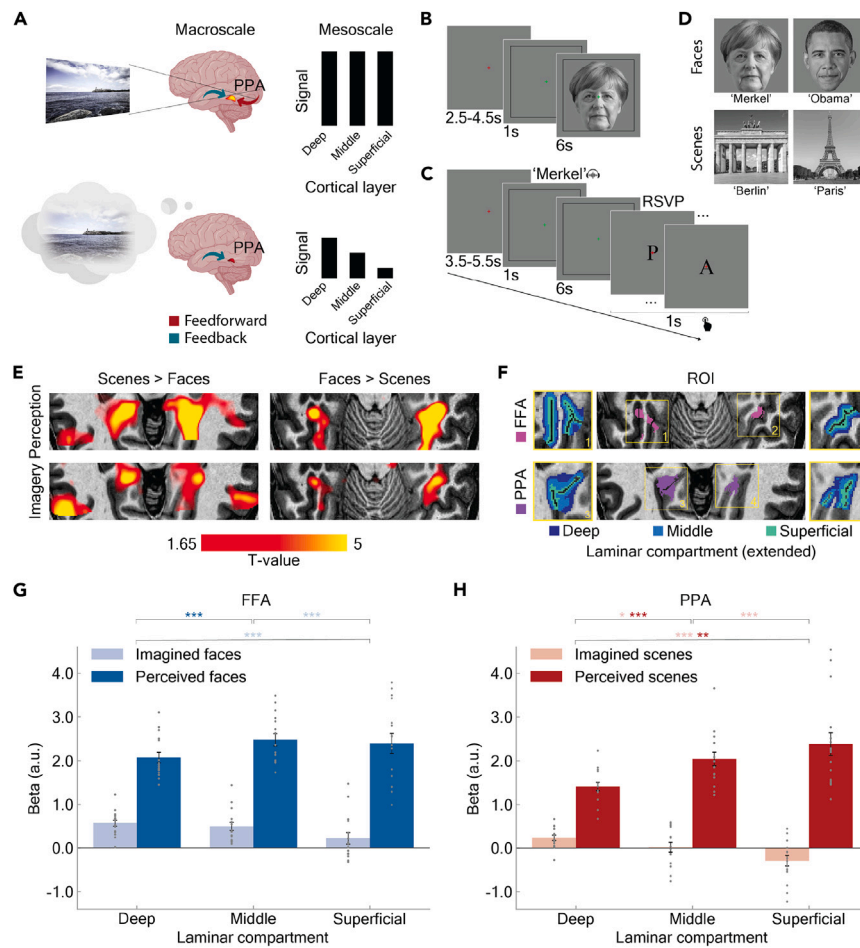


Figure 1. Hypothesis, experimental design, and response at the macro- and mesoscale of cortical organization in high-level ventral visual cortex

(A) Hypothesis and predictions. We hypothesized that the difference between perception and imagery emerges at the mesoscale level of laminar organization. We predicted stable activations across the full cortical ribbon in perception and a selective activation of the deep layers in imagery.

(B and C) Experimental design. After a baseline interval of variable duration subjects either (B) were visually presented with or (C) were cued to imagine one of the familiar faces or scenes (here: Angela Merkel). For imagery trials a rapid serial visual presentation task followed in which subjects were asked to respond to a target letter (here "A").

(D) Stimulus set. The stimulus set consisted of two famous faces and scenes each.

(E) Neural responses at the macroscale level of cortical organization (overlayed on a T1 anatomical image for a representative subject). We observed overlapping responses during perception and imagery for faces and scenes in the fusiform and the parahippocampal cortices respectively.

(F) Laminar compartments (superimposed on an individual T1-weighted image for a representative subject). We defined three laminar compartments (deep, middle, and superficial) following the equi-volume principle along the fusiform and parahippocampal cortices.

(G and H) Cortical responses at the mesoscale of laminar organization in high-level ventral visual cortex to the preferred stimulus object (averaged across participants). Laminar response to perceived and imagined faces and scenes in (G) FFA and (H) PPA. Error bars represent the standard error of the mean; asterisks denote significant *post hoc* pairwise tests: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Gray dots are single subject values.

RESULTS

We recorded sub-millimeter resolution gradient-echo blood oxygenation level-dependent (GE-BOLD) signals with 7 tesla MRI in high-level ventral visual cortex while subjects ($N = 16$) either viewed (Figure 1B) or visually imagined (Figure 1C) famous faces and scenes (Figure 1D). We estimated the cortical responses to perceived and imagined objects using a standard general linear model (GLM).

Responses at the macroscale of cortical organization

We observed activations at the macroscale of cortical organization across the brain in both perception (Figure S1) and imagery (Figure S2). While the activity in ventral visual cortex was weaker during imagery compared to perception, we observed additional ventro-lateral and centro-dorsal responses in a few instances. This could stem from the higher proportion of imagery trials in relation to perception, or potentially

from other confounding sources resulting from the differences between the two experimental conditions (please refer to the [limitations of the study](#) section for a more detailed discussion).

As expected, both perception and imagery of faces and scenes preferentially activated the parahippocampal cortex and to a lesser extent the fusiform cortex, respectively (Figure 1E), across subjects (Figure S3 and Table S1). This warranted further investigation at the mesoscale of cortical responses in the FFA and PPA to their corresponding preferred stimuli (i.e., faces and scenes). This included segmenting the cortical sheet in FFA and PPA into three laminar compartments following the equi-volume principle: deep, middle, and superficial (Figure 1F). We corrected for the draining vein bias of GE-BOLD by applying a leakage model and averaged voxel activity estimates by laminar compartment. We conducted 2×3 repeated measures ANOVAs separately for FFA (Figure 1G) and PPA (Figure 1H) with stimulus mode (perception, imagery) and laminar compartment (deep, middle, superficial) as within-subject factors.

Responses at the mesoscale of cortical organization

There was an expected main effect of stimulus mode (FFA: $F_{(1,15)} = 243.11$, $p_{corr} = 1.1 \times 10^{-10}$; PPA: $F_{(2,30)} = 160.58$, $p = 2.0 \times 10^{-9}$), mirroring the typically stronger activation observed for perception than for imagery.^{22,23} There was no main effect of laminar compartment (FFA: $F_{(1,15)} = 1.59$, $p_{corr} = 0.22$; PPA: $F_{(1,15)} = 2.48$, $p = 0.1$) but, importantly, a significant interaction effect between laminar compartment and stimulus mode (FFA: $F_{(2,30)} = 6.43$, $p_{corr} = 0.018$; PPA: $F_{(2,30)} = 34.81$, $p = 1.51 \times 10^{-8}$). This results pattern was present also without superficial bias deconvolution in FFA (Figure S4A) and PPA (Figure S4B) and was independent of region of interest (ROI) sizes in a wide range (Figure S4C). This robustness mitigates interpretative risks for BOLD signals and their relationship to the underlying neural activity, in particular in cortical layers. Furthermore, the results were also qualitatively equivalent for a voxel sampling approach with full percentage overlapping of columnar compartments across cortical depth in FFA (Figure S5A) and PPA (Figure S5B). Together our results suggest a differentiable pattern of activity for imagery and perception across cortical depth in both FFA and PPA.

To investigate the interaction effect, we used pairwise *post hoc* tests of between-compartment differences for perception and imagery. During perception (Figure 1G and 1H, dark bars) activity was either comparable (FFA_{deep} vs. FFA_{middle} $p_{corr} = 4.4 \times 10^{-3}$; FFA_{deep} vs. FFA_{superficial} $p_{corr} = 0.46$; FFA_{middle} vs. FFA_{superficial} $p_{corr} = 0.67$) or significantly increasing toward the cortical surface (PPA_{deep} vs. PPA_{middle} $p_{corr} = 9.4 \times 10^{-4}$; PPA_{deep} vs. PPA_{superficial} $p_{corr} = 1.6 \times 10^{-4}$; PPA_{middle} vs. PPA_{superficial} $p_{corr} = 0.09$), consistent with feedforward and feedback information across layers^{3,24} with a potential residual superficial bias, greatly reduced in FFA as compared to PPA. In contrast, during imagery (Figure 1G and 1H, light bars), as hypothesized, activity decreased from the deep through the middle to the superficial compartment in both FFA and PPA (FFA_{deep} vs. FFA_{middle} $p_{corr} = 0.39$; FFA_{deep} vs. FFA_{superficial} $p_{corr} = 4.4 \times 10^{-3}$; FFA_{middle} vs. FFA_{superficial} $p_{corr} = 4.2 \times 10^{-3}$; PPA, all $p_{corr} < 0.05$, for statistical detail see figure caption).

Together, the results in FFA and PPA independently suggest that imagery and perception activate high-level ventral visual cortex differently at the mesoscale level of cortical layers.

DISCUSSION

Our findings have two key implications. First, they provide insight into the neural basis of the phenomenological similarities and differences of visual imagery and perception. The similarity is mediated by the activation of similar cortical macroscale regions representing complex visual contents.^{1,25} The dissimilarity in contrast is mediated by the distinctive mesoscale laminar activation, with imagery characterized by deep-layer activation and perception by stable activation across all layers.⁹ This multiscale organization of the visual cortex offers a parsimonious explanation of how object imagery and perception may co-occur while having a specific signature. In turn, it predicts that, in atypical phenomenological conditions like hallucinations, layer segregation of feedback signals may be disturbed.

Second, our findings shed light on how feedforward and feedback neural communication orchestrates object vision. Previous research efforts dissociating information flow by cortical layers focused on low-level visual areas^{7,12,17,26} or prefrontal cortex.²⁷ Here we show that feedback and feedforward processing has a distinct layer-specific profile in high-level ventral visual cortex, the locus of sensory evidence of perceptual decision making about complex visual stimuli.²⁸ Together our results suggest a relay function within high-level ventral visual cortex during mental imagery, closing a crucial gap in tracking feedback information across the processing cascade.^{2,29}

We observed feedback-related activity in deep, but not superficial, layers. Previous studies offer a mixed picture with feedback activity in both deep and superficial layers,^{13,26} or either of them.^{9,10,12} A likely explanation is that distinct modes of feedback exist,^{9,15} with externally and internally generated feedback targeting primarily superficial and deep layers, respectively. Consistent with this idea and our results, previous studies in high-level ventral visual cortex observed effects of memory-retrieval in deep layers,¹⁶ while task demands modulated superficial layer activity.¹¹ Our work paves the way to future studies that simultaneously modulate externally and internally generated feedback signals in high-level ventral visual cortex.

Limitations of the study

A potential study limitation arises from the design differences between the perception and imagery conditions, particularly due to the confounding influence of the auditory cue and the rapid serial visual presentation (RSVP) task. Although we regressed their effect out in our GLM analysis, nonlinearity and multicollinearity dynamics might still exist and contribute to the observed stronger responses in imagery beyond the ROIs. However, the interpretation remains challenging given the typically observed differences in signal-to-noise ratios typically observed

between imagery and perception³⁰ which we aimed to counteract by having four times as many trials in the imagery than in the perception condition.

An alternative source of feedback modulation, different from feedback induced by visual imagery, cannot be ruled out. For example, auditory stimulation could have led to top-down cross-modal deactivation.³¹ Similarly, the RSVP task could have led to top-down preparatory attention^{32,33} or top-down task-demand effect.³⁴ However, beyond the primary visual cortex, both attention and task-demand effects have been previously observed in superficial laminar compartment,^{11,35} contrasting with our findings, which suggest an imagery effect primarily in the deep laminar compartment. Given the limited evidence provided here, future research that fully equates and controls stimulation and task demand in the imagery and perception condition is needed to ultimately rule out the aforementioned alternative explanations.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
 - Lead contact
 - Materials availability
 - Data and code availability
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
- [METHOD DETAILS](#)
 - Stimuli
 - Experimental design
 - Procedure
 - MRI acquisition parameters
 - MRI preprocessing
 - Univariate analysis
 - Regions of interest
 - Superficial bias deconvolution and laminar response quantification
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2024.110229>.

ACKNOWLEDGMENTS

This work was funded by the German Research Foundation (DFG, CI241/1-1, CI241/3-1, and INST 272/297-1 to R.M.C.) and the European Research Council (ERC, 803370 to R.M.C.) and supported by the Einstein Center for Neurosciences Berlin (to T.C.). We thank Maya Jastrzebowska, Maxi Becker, and Johannes Singer for comments on the manuscript. Computing resources were provided by the high-performance computing facilities at ZEDAT, Freie Universität Berlin.

AUTHOR CONTRIBUTIONS

Conceptualization, T.C. and R.M.C.; methodology T.C., R.M.C., and P.I.; software, T.C., P.I., and D.C.; formal analysis and investigation, T.C.; resources, N.W. and R.T.; writing – original draft, T.C. and R.M.C.; writing – review and editing, all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 5, 2023

Revised: January 26, 2024

Accepted: June 6, 2024

Published: June 8, 2024

REFERENCES

- O'Craven, K.M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023. <https://doi.org/10.1162/08989290051137549>.
- Mechelli, A., Price, C.J., Friston, K.J., and Ishai, A. (2004). Where Bottom-up Meets Top-down: Neuronal Interactions during Perception and Imagery. *Cereb. Cortex* 14, 1256–1265. <https://doi.org/10.1093/cercor/bhh087>.
- Larkum, M.E., Zhu, J.J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398, 338–341. <https://doi.org/10.1038/18686>.
- Aru, J., Suzuki, M., and Larkum, M.E. (2020). Cellular Mechanisms of Conscious Processing. *Trends Cogn. Sci.* 24, 814–825. <https://doi.org/10.1016/j.tics.2020.07.006>.
- Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M.A.J., and Friston, K. (2017). Distinct Top-down and Bottom-up Brain Connectivity During Visual Perception and Imagery. *Sci. Rep.* 7, 5677. <https://doi.org/10.1038/s41598-017-05888-8>.
- de Hollander, G., van der Zwaag, W., Qian, C., Zhang, P., and Knapen, T. (2021). Ultra-high field fMRI reveals origins of feedforward and feedback activity within laminae of human ocular dominance columns. *Neuroimage* 228, 117683. <https://doi.org/10.1016/j.neuroimage.2020.117683>.
- Lawrence, S.J., Norris, D.G., and de Lange, F.P. (2019). Dissociable laminar profiles of concurrent bottom-up and top-down modulation in the human visual cortex. *Elife* 8, e44422. <https://doi.org/10.7554/eLife.44422>.
- Qian, C., Chen, Z., Hollander, G. de, Knapen, T., Zhang, Z., He, S., and Zhang, P. (2023). Hierarchical and fine-scale mechanisms of binocular rivalry for conscious perception. Preprint at bioRxiv. <https://doi.org/10.1101/2023.02.11.528110>.
- Bergmann, J., Petro, L.S., Abbatecola, C., Li, M.S., Morgan, A.T., and Muckli, L. (2024). Cortical depth profiles in primary visual cortex for illusory and imaginary experiences. *Nat. Commun.* 15, 1002. <https://doi.org/10.1038/s41467-024-45065-w>.
- Aitken, F., Menelaou, G., Warrington, O., Koolschijn, R.S., Corbin, N., Callaghan, M.F., and Kok, P. (2020). Prior expectations evoke stimulus-specific activity in the deep layers of the primary visual cortex. *PLoS Biol.* 18, e3001023. <https://doi.org/10.1371/journal.pbio.3001023>.
- Dowdle, L., Ghose, G., Moeller, S., Ugurbil, K., Yacoub, E., and Vizioli, L. (2023). Characterizing top-down microcircuitry of complex human behavior across different levels of the visual hierarchy. Preprint at bioRxiv. <https://doi.org/10.1101/2022.12.03.518973>.
- Muckli, L., De Martino, F., Vizioli, L., Petro, L.S., Smith, F.W., Ugurbil, K., Goebel, R., and Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Curr. Biol.* 25, 2690–2695. <https://doi.org/10.1016/j.cub.2015.08.057>.
- Lawrence, S.J.D., van Mourik, T., Kok, P., Koopmans, P.J., Norris, D.G., and de Lange, F.P. (2018). Laminar Organization of Working Memory Signals in Human Visual Cortex. *Curr. Biol.* 28, 3435–3440.e4. <https://doi.org/10.1016/j.cub.2018.08.043>.
- Van Kerkoerle, T., Self, M.W., and Roelfsema, P.R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nat. Commun.* 8, 13804. <https://doi.org/10.1038/ncomms13804>.
- Koenig-Robert, R., and Pearson, J. (2021). Why do imagery and perception look and feel so different? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376, 20190703. <https://doi.org/10.1098/rstb.2019.0703>.
- Koyano, K.W., Takeda, M., Matsui, T., Hirabayashi, T., Ohashi, Y., and Miyashita, Y. (2016). Laminar Module Cascade from Layer 5 to 6 Implementing Cue-to-Target Conversion for Object Memory Retrieval in the Primate Temporal Cortex. *Neuron* 92, 518–529. <https://doi.org/10.1016/j.neuron.2016.09.024>.
- Kok, P., Bains, L.J., van Mourik, T., Norris, D.G., and de Lange, F.P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Curr. Biol.* 26, 371–376. <https://doi.org/10.1016/j.cub.2015.12.038>.
- Rockland, K.S., and Pandya, D.N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* 179, 3–20. [https://doi.org/10.1016/0006-8993\(79\)90485-2](https://doi.org/10.1016/0006-8993(79)90485-2).
- Salin, P.A., and Bullier, J. (1995). Corticocortical connections in the visual system: structure and function. *Physiol. Rev.* 75, 107–154. <https://doi.org/10.1152/physrev.1995.75.1.107>.
- van Kerkoerle, T., Self, M.W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., and Roelfsema, P.R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. USA* 111, 14332–14341. <https://doi.org/10.1073/pnas.1402773111>.
- Barzegaran, E., and Plomp, G. (2022). Four concurrent feedforward and feedback networks with different roles in the visual cortical hierarchy. *PLoS Biol.* 20, e3001534. <https://doi.org/10.1371/journal.pbio.3001534>.
- Tullo, M.G., Almgren, H., Van de Steen, F., Boccia, M., Bencivenga, F., and Galati, G. (2023). Preferential signal pathways during the perception and imagery of familiar scenes: An effective connectivity study. *Hum. Brain Mapp.* 44, 3954–3971. <https://doi.org/10.1002/hbm.26313>.
- Lee, S.-H., Kravitz, D.J., and Baker, C.I. (2012). Disentangling visual imagery and perception of real-world objects. *Neuroimage* 59, 4064–4073. <https://doi.org/10.1016/j.neuroimage.2011.10.055>.
- Klink, P.C., Dagnino, B., Gariel-Mathis, M.-A., and Roelfsema, P.R. (2017). Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. *Neuron* 95, 209–220.e3. <https://doi.org/10.1016/j.neuron.2017.05.033>.
- Ganis, G., Thompson, W.L., and Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Brain Res. Cogn. Brain Res.* 20, 226–241. <https://doi.org/10.1016/j.cogbrainres.2004.02.012>.
- Haarsma, J., Deveci, N., Corbin, N., Callaghan, M.F., and Kok, P. (2023). Expectation Cues and False Percepts Generate Stimulus-Specific Activity in Distinct Layers of the Early Visual Cortex. *J. Neurosci.* 43, 7946–7957. <https://doi.org/10.1523/JNEUROSCI.0998-23.2023>.
- Finn, E.S., Huber, L., Jangraw, D.C., Molfese, P.J., and Bandettini, P.A. (2019). Layer-dependent activity in human prefrontal cortex during working memory. *Nat. Neurosci.* 22, 1687–1695. <https://doi.org/10.1038/s41593-019-0487-z>.
- McKeeff, T.J., Olman, F., and Tong, F. (2007). The Timing of Perceptual Decisions for Ambiguous Face Stimuli in the Human Ventral Visual Cortex. *Cereb. Cortex* 17, 669–678. <https://doi.org/10.1093/cercor/bkh015>.
- Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nat. Rev. Neurosci.* 20, 624–634. <https://doi.org/10.1038/s41583-019-0202-9>.
- Naselaris, T., Olman, C.A., Stansbury, D.E., Ugurbil, K., and Gallant, J.L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage* 105, 215–228. <https://doi.org/10.1016/j.neuroimage.2014.10.018>.
- Gau, R., Bazin, P.-L., Trampel, R., Turner, R., and Noppeney, U. (2020). Resolving multisensory and attentional influences across cortical depth in sensory cortices. *Elife* 9, e46856. <https://doi.org/10.7554/eLife.46856>.
- Stokes, M., Thompson, R., Nobre, A.C., and Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci. USA* 106, 19569–19574. <https://doi.org/10.1073/pnas.0905306106>.
- Battistoni, E., Stein, T., and Peelen, M.V. (2017). Preparatory attention in visual cortex. *Ann. N. Y. Acad. Sci.* 1396, 92–107. <https://doi.org/10.1111/nyas.13320>.
- Koida, K., and Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nat. Neurosci.* 10, 108–116. <https://doi.org/10.1038/nn1823>.
- Nandy, A.S., Nassi, J.J., and Reynolds, J.H. (2017). Laminar Organization of Attentional Modulation in Macaque Visual Area V4. *Neuron* 93, 235–246. <https://doi.org/10.1016/j.neuron.2016.11.029>.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>.
- Huber, L.R. (Renzo), Poser, B.A., Bandettini, P.A., Arora, K., Wagstyl, K., Cho, S., Goense, J., Nothnagel, N., Morgan, A.T., van den Hurk, J., et al. (2021). LayNii: A software suite for layer-fMRI. *Neuroimage* 237, 118091. <https://doi.org/10.1016/j.neuroimage.2021.118091>.
- Downing, P.E., Chan, A.W.Y., Peelen, M.V., Dods, C.M., and Kanwisher, N. (2006). Domain specificity in visual cortex. *Cereb. Cortex* 16, 1453–1461. <https://doi.org/10.1093/cercor/bhj086>.
- Grill-Spector, K., Knouf, N., and Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7, 555–562. <https://doi.org/10.1038/nn1224>.

40. Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res.* 41, 1409–1422. [https://doi.org/10.1016/S0042-6989\(01\)00073-6](https://doi.org/10.1016/S0042-6989(01)00073-6).
41. Epstein, R., Graham, K.S., and Downing, P.E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37, 865–876. [https://doi.org/10.1016/S0896-6273\(03\)00117-X](https://doi.org/10.1016/S0896-6273(03)00117-X).
42. Amunts, K., Mohlberg, H., Bludau, S., and Zilles, K. (2020). Julich-Brain: A 3D probabilistic atlas of the human brain's cytoarchitecture. *Science* 369, 988–992. <https://doi.org/10.1126/science.abb4588>.
43. Yushkevich, P.A., Pashchinskiy, A., Oguz, I., Mohan, S., Schmitt, J.E., Stein, J.M., Zukić, D., Vicory, J., McCormick, M., Yushkevich, N., et al. (2019). User-Guided Segmentation of Multi-modality Medical Imaging Datasets with ITK-SNAP. *Neuroinformatics* 17, 83–102. <https://doi.org/10.1007/s12021-018-9385-x>.
44. Lüsebrink, F., Sciarra, A., Mattern, H., Yakupov, R., and Speck, O. (2017). T1-weighted in vivo human whole brain MRI dataset with an ultrahigh isotropic resolution of 250 μm . *Sci. Data* 4, 170032. <https://doi.org/10.1038/sdata.2017.32>.
45. Zaretskaya, N., Fischl, B., Reuter, M., Renvall, V., and Polimeni, J.R. (2018). Advantages of cortical surface reconstruction using submillimeter 7 T MEMPRAGE. *Neuroimage* 165, 11–26. <https://doi.org/10.1016/j.neuroimage.2017.09.060>.
46. Waehnert, M.D., Dinse, J., Weiss, M., Streicher, M.N., Waehnert, P., Geyer, S., Turner, R., and Bazin, P.-L. (2014). Anatomically motivated modeling of cortical laminae. *Neuroimage* 93 Pt 2, 210–220. <https://doi.org/10.1016/j.neuroimage.2013.03.078>.
47. Polimeni, J.R., Fischl, B., Greve, D.N., and Wald, L.L. (2010). Laminar analysis of 7T BOLD using an imposed spatial activation pattern in human V1. *Neuroimage* 52, 1334–1346. <https://doi.org/10.1016/j.neuroimage.2010.05.005>.
48. Markuerkiaga, I., Marques, J.P., Gallagher, T.E., and Norris, D.G. (2021). Estimation of laminar BOLD activation profiles using deconvolution with a physiological point spread function. *J. Neurosci. Methods* 353, 109095. <https://doi.org/10.1016/j.jneumeth.2021.109095>.
49. Markuerkiaga, I., Barth, M., and Norris, D.G. (2016). A cortical vascular model for examining the specificity of the laminar BOLD signal. *Neuroimage* 132, 491–498. <https://doi.org/10.1016/j.neuroimage.2016.02.073>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw data	This study	https://osf.io/4xru2/
Software and algorithms		
Code for processing and analyzing the data	This study	https://github.com/carricarte/Face-Scene_7T-MRI
FSL	Smith et al. ³⁶	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation/Linux
LAYNII	Huber et al. ³⁷	https://github.com/layerfMRI/LAYNII

RESOURCE AVAILABILITY

Lead contact

Further information and requests may be directed to and will be fulfilled by the lead contact, Tony Carricarte (tcarricarte@gmail.com).

Materials availability

No materials were generated in this study.

Data and code availability

- Data: All original code has been deposited at Open Science Framework repository and is publicly available. DOIs are listed in the [key resources table](#).
- Code: All original code has been deposited at GitHub and is publicly available. DOIs are listed in the [key resources table](#).
- Additional information: Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Eighteen adult volunteers participated in the study after providing written informed consent. Two subjects were excluded from the final analysis due to technical errors during the scanning session. The analyzed sample was thus $N = 16$ (mean age 27.5 years; age range 18–38 years; 8 female). All subjects had normal or corrected-to-normal vision and no history of neurological disorders. All subjects received a monetary reward at the end of the study. The study was approved by the Ethics Committee of the medical school of the University of Leipzig (Study approval number: 035/22-ek) and was conducted in accordance with the Declaration of Helsinki.

METHOD DETAILS

Stimuli

Our goal was to explore the mesoscale response in high-level ventral visual cortex, a region known for its selective activation to complex visual stimuli.³⁸ To ensure robust activations within this area, we specifically chose faces and scenes based on previous studies.^{39–41} The stimuli consisted of two gray-scale images of highly famous faces and places each – the former German chancellor Angela Merkel, the former US president Barack Obama, the Brandenburg Gate and the Eiffel Tower (Figure 1D). We chose these stimuli based on a brief survey ($N=7$) aimed at rating the ease with which famous faces and scenes could be imagined. Given that the signal to noise ratio during imagery is very low, we selected the top-rated stimuli that participants could effectively visualize.

Experimental design

We conducted a main experiment and a standard localizer experiment. The main experiment included between nine and twelve runs, with each run lasting approximately 256 s. Each run contained imagery and perception trials with all four object stimuli interleaved in pseudo-randomized order. Given that imagery elicits weaker activations compared to perception,²³ each run consisted of four perception trials and sixteen imagery trials. We used this proportion in order to increase effective contrast-to-noise ratio with respect to the imagery condition.

Each perception trial started with a baseline interval (jittered duration 2.5–4.5 s in 0.5 s steps) during which a red fixation cross was centrally displayed on a gray background (Figure 1B). This was followed by a 1-s period during which a black $5^\circ \times 5^\circ$ square frame appeared centrally and the color of the fixation cross turned green to signal the upcoming task. Next, one of the four stimuli was displayed intermittently within the square frame for a total duration of 6 s. Each intermittent display time lasted 250 ms on and 250 ms off.

Each imagery trial started with a jittered baseline interval (3.5–5.5 s in 0.5 s steps) during which a red fixation cross was centrally displayed on a gray background. This was followed by a 1-s period during which a black $5^\circ \times 5^\circ$ square frame appeared centrally and the color of the fixation cross turned green to signal the upcoming task. At the onset of this period of a given imagery trial, a voice recording ('Merkel', 'Obama', 'Berlin', 'Paris') was briefly played to indicate the desired content of the upcoming imagery task. The voice recording presentation signaled an imagery trial or a perception trial otherwise. The imagery task lasted 6 s and consisted of visually imagining the auditorily cued stimulus (Figure 1C). Subjects were instructed to imagine the relevant target object within the displayed frame, all while keeping their eyes open. To control the duration of the imagery task, we asked subjects to perform a letter detection task at the end of the trial. Subjects reported the appearance of the character 'A', with fifty percent presentation probability, as soon as it was detected from a rapid serial visual presentation of 1-s duration. The baseline variation between perception and imagery trials was due to the inclusion of an extra 1-sec response period in the RSVP task, which was visually equivalent to the baseline interval. Subjects were instructed to fixate on the central fixation cross throughout the scanning session.

We also included a standard localizer experiment after the main experiment to locate the regions of interest FFA and PPA for each subject. The localizer consisted of a single run of 500 s, comprising six randomly presented blocks each of faces, places, objects, scrambled objects and blank background. Each block lasted 16 s and was composed of twenty different stimuli, displayed for 500 ms with a 300-ms interstimulus interval in a random order. During the run, subjects performed a one-back repetition task in which they had to indicate when a currently presented image was the same as the one that immediately preceded it.

Procedure

MRI data were acquired at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany. Each subject completed two scanning sessions on two separate days. During the first scanning session, we acquired a T1-weighted anatomical image and two short FFA and PPA localizers. We used these short online localizers to guide the positioning of the acquisition slab on the second scanning day. During the second scanning session, we conducted the main and the localizer experiments. To ensure that subjects were familiar with the main task and the stimuli set (Figure 1B), we provided them with verbal and written instructions, including the stimuli set, at least one day before the second scanning session. Additionally, subjects completed a short training of less than twenty trials of the main experiment on the second day before entering the scanner.

MRI acquisition parameters

We acquired MR images on a Siemens Magnetom Terra 7T whole-body system (Siemens Healthcare, Erlangen, Germany) with a single-channel-transmit and a 32-channel radio-frequency (RF) receive head coil (Nova Medical Inc, Wilmington, USA). We acquired the functional data using a 2D Gradient-echo (GE) echo planar imaging (EPI) sequence (voxel size = 0.8 mm isotropic resolution, TE/TR = 25/3000 ms, in-plane field of view (FoV) $148 \times 148 \text{ mm}^2$, 48 coronal slices, flip angle = 70° , echo spacing = 1.01 ms, GRAPPA factor = 3, partial Fourier = 6/8, phase encoding direction head-foot). We recorded anatomical data using an MP2RAGE sequence (voxel size = 0.7 mm isotropic resolution, TE/TR = 2.01/5000 ms, in-plane FoV $224 \times 224 \text{ mm}$, GRAPPA factor = 2) yielding two inversion contrasts (T11 = 900 ms, flip angle = 5° ; T12 = 2750 ms, flip angle = 3°). The two inversion contrasts were combined to produce T1-weighted MP2RAGE uniform (UNI) images with high contrast to noise ratio.

MRI preprocessing

Functional data: We spatially realigned the functional images to the first volume of the first run using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>). To ensure that volumes were well aligned, we computed the mean intensity correlation between each pair of runs. Spatial correlation between all pairs of runs were above 0.9 and included in further processing.

To co-register anatomical and functional images, we initially referred to the Julich-Brain atlas⁴² to identify the approximate locations of the fusiform and parahippocampal regions, and then we outlined these regions in ventral temporal cortex of both hemispheres on the individual subject's native space (from here on referred to as manual masks) using ITK-SNAP with the 3D paintbrush tool (v.3.8).⁴³ We then registered the motion-corrected images to the anatomical volume using the Advanced Normalization Tools (ANTs) software package (<http://stnava.github.io/ANTs/>), applying linear (rigid) and nonlinear (SyN) transformations within the manual mask defined previously. We visually inspected the fixed and registered volumes in ITK-SNAP for each subject. If the volumes were not correctly registered within the region of the manual mask, we repeated the registration using a different combination of parameters (e.g., omitting rigid transform, including affine transform and using skull-stripped volumes) until an accurate alignment was achieved. Next, we resampled the functional images to the anatomical reference using ANTs transform function with five order b-spline interpolation. Finally, we spatially smoothed the functional localizer images using a 6-mm full width at half maximum (FWHM) Gaussian kernel. We smoothed the functional images from the main experiment only for visualization purposes but not for the layer-specific analysis to preserve spatial specificity.

Anatomical data: We corrected the T1-weighted UNI volume for bias field effects using a customized script.⁴⁴ Next, we preprocessed the bias-field-corrected volume using FreeSurfer (<http://surfer.nmr.mgh.harvard.edu/>) with the '-hires' flag⁴⁵ to preserve the submillimeter spatial resolution. However, due to inaccuracies in the automatic brain tissue segmentation, we used the conformed anatomical image output to manually delineate the boundaries between gray matter and white matter and cerebrospinal fluid around in functionally defined regions of interest (see below). We performed the segmentation in FSLeves, i.e., the image viewer of FSL (v.6.0.5),³⁶ following the procedure as described in (<https://layerfmri.com/2018/03/11/quick-layering/>). Next, we segmented the cortical ribbon into laminar and columnar

compartments using LAYNII (v2.2.1).³⁷ In detail, applying the equi-volume principle,⁴⁶ we segmented the gray matter into three laminar compartments: deep, middle and superficial (Figure 1C). Here, we used the term laminar compartments to refer to the depth-dependent compartments along the cortical ribbon, distinct from the actual anatomical layers found in cortex. Additionally, we segmented the gray matter into columnar compartments within the manual masks (range: 890-1980).

Univariate analysis

We conducted all analyses separately in each subject's native anatomical space. We ran two separate General Linear Model (GLM) analyses in SPM12 for each pre-processed functional dataset, one for the main experiment within the manual mask, and another for the localizer experiment. Specifically, we constructed a regressor for the baseline and each experimental condition, perceived and imagined faces and scenes (main experiment), and faces and scenes (localizer experiment) by convolving a boxcar function representing the onsets and durations of the corresponding trials with the canonical (2 Gamma) hemodynamic response function (HRF). We incorporated the motion estimates along with the auditory cue and RSVP task into the model as nuisance regressors in order to remove their effects.

Regions of interest

We defined regions-of-interest (ROIs) based on the GE-BOLD functional localizer run. Superficial draining veins strongly contribute to the GE-BOLD signal, biasing the sampling of voxels towards those closer to the pial surface. To reduce this sampling bias, we defined FFA and PPA compartment-specific ROIs by ranking the voxels within each laminar compartment according to the corresponding T-statistic and retaining only the top 500 voxels. In detail, for each compartment of the FFA, we ranked the voxels according to the faces > places localizer contrast T-statistic. For each compartment of the PPA, we ranked the voxels according to the places > faces localizer contrast T-statistic. The activation maps were further intersected with the previously drawn manual mask to anatomically restrict them to likely locations of FFA and PPA. ROIs were separately defined for the left and right hemispheres and then merged into a single bilateral ROI, resulting in one bilateral FFA and one bilateral PPA per subject, each one containing 1000 voxels in total in each laminar compartment. This yielded on average around 50% overlap of voxels within columnar compartments across cortical depth (Table S2).

Superficial bias deconvolution and laminar response quantification

The GE-BOLD signal is strongly affected by locally nonspecific responses from macrovasculature,⁴⁷ compromising the estimation of the laminar response. To address this, we applied a deconvolution model,⁴⁸ assuming that the response in a given layer is composed of the activity in that layer plus a nonspecific leaked response coming from the layers beneath. The model works by deconvolving the measured BOLD signals with a physiological Point Spread Function (PSF).⁴⁹ The PSF characterizes the BOLD signal leakage from the layer of activation to downstream layers.

Initially, we estimated the amplitude of the low frequency (ALF) on the localizer run using FSL. We then calculated the local activity specific to each laminar compartment using LAYNII function LN2_DEVEIN. For this, we included the ALF estimates together with the previously calculated laminar and columnar compartments. We set the lambda parameter to 0.25, i.e., assuming a moderate baseline cerebral blood flow. Finally, for each ROI, we calculated the laminar response to the preferred stimuli, with and without superficial bias deconvolution, by averaging the beta values across all voxels within the given laminar compartment.

QUANTIFICATION AND STATISTICAL ANALYSIS

To quantify the differences between the responses to perceived and imagined preferred stimuli across the laminar compartments, we ran two-way repeated-measures analyses of variance (rm-ANOVA), separately for the FFA and PPA ROIs. We treated stimulus mode (perception and imagery) and laminar compartment (deep, middle and superficial) as within-subject factors. Before running the rm-ANOVA, we tested for normality and sphericity by running the Shapiro-Wilk and Mauchly tests, respectively. In case of violation of the sphericity assumption, we used the Greenhouse-Geisser correction. We applied the Holm-Bonferroni correction for post-hoc test adjustments to account for multiple comparisons. All statistical tests were conducted with a significance level $\alpha = 0.05$.