Trial-by-trial inter-areal interactions in visual cortex in the presence or absence of visual stimulation

Dianna Hidalgo[1*†], Giorgia Dellaferrera[2†], Will Xiao[1], Maria Papadopouli[3,4], Stelios Smirnakis[1,5], Gabriel Kreiman[2*]

***For correspondence:** gabriel.kreiman@tch.harvard.edu (GK); diannahidalgo@g.harvard.edu (DH)

[†]These authors contributed equally to this work

[1]Department of Neurobiology, Harvard Medical School, Boston, MA, USA.; [2]Children's Hospital, Harvard Medical School, Boston, MA, USA.; [3]Department of Computer Science, University of Crete, Heraklion, Greece; [4]Institute of Computer Science, FORTH, Heraklion, Greece; [5]Department of Neurology, Brigham and Women's Hospital, Boston,MA, USA

Abstract State-of-the-art computational models of vision largely focus on fitting trial-averaged spike counts to visual stimuli using overparameterized neural networks. However, a computational model of the visual cortex should predict the dynamic responses of neurons in single trials across different experimental conditions. In this study, we investigated trial-by-trial inter-areal interactions in the visual cortex by predicting neuronal activity in one area based on activity in another, distinguishing between stimulus-driven and non-stimulus-driven shared variability. We analyzed two datasets: calcium imaging from mouse V1 layers 2/3 and 4, and extracellular neurophysiological recordings from macaque V1 and V4. Our results show that neuronal activity can be predicted bidirectionally between L2/3 and L4 in mice, and between V1 and V4 in macaques, with the latter interaction exhibiting directional asymmetry. The predictability of neuronal responses varied with the type of visual stimulus, yet responses could also be predicted in the absence of visual stimulation. In mice, we observed a bimodal distribution of neurons, with some neurons primarily driven by visual inputs and others showing predictable activity during spontaneous activity despite lacking consistent visually evoked responses. Predictability also depended on intrinsic neuronal properties, receptive field overlap, and the relative timing of activity across areas. Our findings highlight the presence of both stimulus- and non-stimulus-related components in interactions between visual areas across diverse contexts and underscore the importance of non-visual shared variability between visual regions in both mice and macaques.

Introduction

To predict the activity of neurons in the visual cortex, multiple studies have focused on correlating external stimuli with trial-averaged responses (***Hubel and Wiesel, 1962***; ***Pasupathy et al., 2020***). Between the stimulus and cortical neurons, there is a complex signal processing cascade involving multiple processing stages. Therefore, computational models of visual processing typically gloss over most of the relevant biological machinery in an attempt to fit average firing rates from images (***Serre et al., 2007a***; ***Yamins et al., 2014***). A mechanistic understanding of the factors that govern firing in the visual cortex requires models that can capture the trial-by-trial transformations across those processing stages. Moreover, neurons throughout the cortex fire "spontaneously" in the

absence of any visual input. Thus, by definition, any model predicting neuronal activity that is exclusively dependent on visual stimulation does not account for such fluctuations. Previous studies in mice have revealed significant non-visual influences in neuronal activity in cortex, even in V1, partly accounted for by movement (***Stringer et al., 2019b***; ***Avitan and Stringer, 2022***; ***Polack et al., 2013***; ***Niell and Stryker, 2010***; ***Dadarlat and Stryker, 2017***). These observations contrast with a recent macaque study which did not find the same motor-related spontaneous activity in either V1, V2, or V3 (***Talluri et al., 2023***). Nevertheless, variables that are not related to movement, such as attention, expectation, and arousal, also modulate stimulus- and non-stimulus driven neuronal activity in monkeys (***Reynolds and Chelazzi, 2004***; ***Gazzaley et al., 2007***; ***Okazaki et al., 2008***; ***Gilbert and Li, 2013***), potentially adding to the response variability across stimulus repeats and to neuronal activity in the absence of visual stimuli.

Neuronal interactions between visual areas occur in the presence and absence of visual stimuli (***Chen et al., 2022***; ***Stringer et al., 2019b***; ***Wosniack et al., 2021***; ***Ringach, 2009***; ***Avitan and Stringer, 2022***). Therefore, such interactions can and should be studied both as a function of sensory inputs and contextual cues but also in the absence of external stimulation or task demands (***Chacron et al., 2003***; ***Hsu et al., 2004***; ***Ringach, 2009***). A paradigmatic example of inter-area interactions is the series of synaptically-connected laminar (e.g. layer 4 → layer 2/3) and cortical areas (e.g. V1→V2→V4→IT) within the ventral visual stream (***Lee et al., 2016***; ***Felleman and Van Essen, 1991***; ***Markov et al., 2014***; ***Douglas and Martin, 2004***; ***Wang and Burkhalter, 2007***; ***Consortium et al., 2021***). Due to feedforward, feedback, and horizontal connections in the ventral visual stream, the inter-areal interactions could represent shared visual and non-visual reliable information. Several studies examined *in vivo* interactions between visual areas in mice and macaques, focusing on the entire population level (***Semedo et al., 2019***, ***2022***; ***Tang et al., 2023***; ***Morales-Gregorio et al., 2024***), trial-averaged responses removing transient fluctuations (***Semedo et al., 2019***), neuronal activity in response to only one image presentation (***Semedo et al., 2019***, ***2022***), or in the absence of any stimulus (***Morales-Gregorio et al., 2024***). Here we investigated interactions between areas in single trials at the level of cortical layers or brain areas across different stimulus types or in the absence of visual stimulation, across different species, and across different recording techniques and temporal resolutions. We focused on multiple simultaneously recorded areas of the ventral visual stream to assess the stimulus- and non-stimulus-driven variability shared between cortical subnetworks. We found that it is possible to reciprocally predict neuronal activity, both during visual stimulation but also during spontaneous activity, and that this predictability depends on the intrinsic properties of each neuron, the degree of receptive field overlap, and the relative timing of activity across areas.

Results

Layer 4 activity predicts layer 2/3 activity and V1 activity predicts V4 activity in single trials

We studied neuronal activity from two open datasets: mouse neurons in V1 layer 4 and layers 2/3 (L4 and L2/3; calcium imaging; ***Figure 1***A) (***Stringer et al., 2019a***), and macaque multiunit sites in areas V1 and V4 (extracellular electrophysiology; ***Figure 1***B) (***Chen et al., 2022***). The mouse neuronal recordings we used for this experiment were based on approx. 5,500 per mouse (n=4, ***Table 1***) responding to visual stimuli (drifting gratings or static natural black and white images; total of 7 recording days), in addition to "spontaneous" activity during approximately 30 minutes of gray/black screen presentation on 6 of the 7 recording days. The macaque recordings were based on 688 out of the 1,024 channels (n=1, ***Table 2***) responding to visual stimuli (full-size static checkerboard image, small and thin bar slow-moving in a small clockwise square direction; large and thick bar fast-moving in a big clockwise square direction; total of 5 recording days) in addition to spontaneous activity during gray screen presentation in all recording days. There was also a lights-off condition, where the head-fixed monkey was free to open or close eyes for approximately 25 min-

utes in 3 our of the 5 recording days. We omitted some of the 1,024 channels with signal-to-noise ratio of less than 2, or that were considered "spurious" by the authors of the ope dataset (***Chen et al., 2022***).

We examined two types of interactions between areas: inter-laminar (***Figure 1***C; mouse V1) and inter-cortical (***Figure 1***D; macaque). We used linear ridge regression to predict neuronal activity in one area from activity in the other area in single trials (***Figure 1***E) (***Semedo et al., 2019***). Performance was evaluated using cross-validation over trials and quantified as squared Pearson's r (hereafter, "explained variance" or EV, Methods). ***Figure 2***A shows sample neuronal activity from three example mouse V1 L2/3 cells during image presentation (black traces). Overlaid, the figure also shows the predicted neuronal activity (red). The predicted neuronal activity is shown as a function of the actual activity in response to every image presentation for the same example cells in ***Figure 2***C. The top cell illustrates a case where the predicted activity closely matches the actual actual activity ($EV = 0.67$), the middle cell shows a typical case ($EV = 0.39$), and the bottom cell illustrates a case where the predictions deviated from the actual neuronal activity ($EV = 0.07$). We focused on neurons deemed "visually responsive" (~17% of total L2/3 neurons; ***Table 1***, Methods, see results for all neurons in ***Figure Supplement 1***). The ridge regression model predicted single-trial L2/3 activity from L4 activity across both types of visual stimuli with an average EV of 0.28 ± 0.16 (mean ± stdev. across neurons, ***Figure 2***E) whereas the shuffle control mean EV was 0.004 ± 0.002 (see results for individual mice in ***Figure Supplement 1***).

In the macaque, trial-to-trial variations in V4 activity were predicted from V1 activity across the three types of visual stimuli. Example recording sites are shown in ***Figure 2***B, D. The ridge regression model predicted the single-trial responses in V4 activity from V1 activity with an average EV of 0.34 ± 0.15 (***Figure 2***F whereas the shuffle control mean EV was 0.005 ± 0.005. There were few sites that were not visually responsive in macaques; EV results for all sites are shown in ***Figure Supplement 1***).

In sum, it was possible to provide estimates of neuronal activity in single trials in both species, across different layers within primary visual cortex in mice and across different visual cortical areas in monkeys.

Inter-cortical predictions are asymmetrical

In the previous section, we demonstrated the possibility of predicting L2/3 activity from L4 activity and V4 from V1. We asked whether we could also predict neuronal responses in the opposite direction. To directly compare predictability between directions in mouse and macaque, we matched the number of predictors (i.e., number of neurons/sites used to predict activity) and the degree of self-consistency (split-half r values) by randomly subsampling in each layer or cortical region prior to computing the predictability metrics (***Figure 3***A, C, Methods).

In mice, it was possible to predict L4 neuronal activity from the activity of populations of neurons in L2/3 and there was no statistically significant difference between the two directions ($p > 0.05$, hierarchical permutation test, ***Figure 3***B). When using the entire layer populations to predict each other's neural activity (923-2,369 cells in L4, 5,420-7,980 cells in L2/3), L2/3 could predict L4 better than the reverse direction ($p < 0.05$, ***Figure Supplement 2***G).

In macaques, while we could also predict V1 activity from the activity of a population of neurons in V4, when controlling for neuron number and split-half correlation values, the EV fraction in the V1→V4 direction was higher than in the V4→V1 direction ($p < 0.001$, ***Figure 3***D). Even without controlling for the number of predictors or their respective split-half correlation values (627-688 sites in V1, 86-115 sites in V4), we found better predictability in the V1 to V4 direction than the reverse ($p < 0.001$, ***Figure Supplement 2***I).

Predictability of neuronal activity is dependent on the visual stimulus

We evaluated whether the predictability of neuronal activity varied with the type of visual stimulus presented to the animal. In mice, we compared the inter-laminar prediction of neuronal activity of

visually responsive neurons in response to drifting gratings versus natural images (***Figure 4***A). We could predict mouse L4 and L2/3 activity under both stimulus conditions ($p < 0.001$, paired permutation test of prediction vs. shuffled frames prediction). Predictability was higher for drifting gratings than natural images in the L4→L2/3 direction (***Figure 4***B; $p < 0.001$, hierarchical permutation test).

In macaques, we compared inter-cortical predictability of visually responsive site recordings in responses to a slow-moving small thin bar, fast-moving large thick bar, and a full-size checkerboard image (***Figure 4***C). We could predict V1 and V4 activity across all stimulus types ($p < 0.001$, paired permutation test of prediction vs. shuffled frames prediction). The predictability was the highest in both directions for neuronal activity in response to a full field checkerboard images (***Figure 4***D). In the V1→V4 direction, the EV fraction was higher when predicting a slow moving small thin bar compared to a fast moving large thick bar (***Figure 4***D, left), where as the opposite was true for the V4→V1 direction (***Figure 4***D, right).

Neuronal activity could be predicted even during spontaneous activity

Given the dependence on the visual stimulus, we next asked whether it would be possible to predict neuronal responses in the absence of any visual stimulus, during "spontaneous activity". We compared the predictability of stimulus-evoked activity in mice (drifting gratings and natural images) versus the predictability of activity recorded during gray screen presentation. This comparison was conducted in both visually (SNR >2 & split-half r >0.8) and non-visually (SNR <2 & split-half r <0.8) responsive neurons (n=3 mice; mouse MP027 did not undergo 30 min. of gray screen presentation). In visually responsive neurons, there was a significant reduction in EV during gray screen compared to visual stimulus presentation (***Figure 5***A left, $p < 0.001$, hierarchical paired permutation test). In contrast, for non-visually responsive neurons, predictability was higher during the gray screen condition (***Figure 5***A right, $p < 0.001$, hierarchical paired permutation test). Additionally, there was no correlation between neuronal predictability in the responses to visual stimulus presentations and in the response to gray screen presentations in visually responsive neurons (***Figure 5***B) but there was a strong correlation for non-visually responsive neurons (***Figure 5***C). The difference in predictability in the absence of a stimulus could in principle change according to the directionality in inter-laminar interactions. There was no statistically significant difference in the EV fraction between laminar directions (L4→L2/3 vs. L2/3→L4) using the same control population as in ***Figure 3***B (***Figure 5***A-C and ***Figure Supplement 2***H).

In macaques, we focused on visually responsive sites since the majority of the neuronal population was visually responsive (***Figure Supplement 1***D). Additionally, an SNR of less than 2 (one of the requirements to define non-visual neurons in the mouse data) most likely reflects artefactual issues with the electrode recording the multiunit site (***Chen et al., 2022***). We compared inter-areal prediction of stimulus presentation activity (checkerboard images and moving bars), gray screen presentation, and during lights-off. Similar to the conclusions drawn from the mouse data, the predictability of neuronal activity was higher in response to stimulus presentation than to gray screen presentations (***Figure 5***D for checkerboard presentations, ***Figure Supplement 4***D for moving bars; $p < 0.001$, paired permutation test). However, the EV fraction in the lights-off condition was significantly higher than during the stimulus presentations in both directions. Eye closure and sleep can induce global oscillations (***Hohaia et al., 2022***) and therefore may correlate neuronal activity, causing an increase in predictability. To test this idea, we further separated the lights-off neuronal activity into periods where the macaque's eyes were open or closed. The EV was higher than stimulus presentation activity only during the eyes-closed period (***Figure 5***D). Unlike the mouse, macaque correlation of visual predictability between stimulus presentation and spontaneous activity was high across all types of spontaneous conditions (***Figure 5***E, ***Figure Supplement 4***C). When assessing the inter-cortical prediction directionality during spontaneous conditions, we found the same asymmetrical relationship as in ***Figure 3***, where V1→V4 EV fraction was significantly higher than V4→V1 prediction in both gray screen ($p < 0.01$, permutation test) and lights-off ($p < 0.001$, permutation test) conditions (***Figure Supplement 2***G).

Receptive field overlap and neuronal response properties impact predictability

We investigated which neuronal properties are related to the ability to predict responses by comparing EV and key indicators of neuronal response reliability and receptive field properties, in both visually- and non-visually responsive neurons, during either visual presentations or spontaneous conditions. First, we considered the following properties: (i) max r^2 value (i.e., maximum squared correlation between each neuron in the predictor population and the predicted neuron), (ii) 1-vs-rest r^2 (the squared correlation of one neuron's activity across all stimuli with the mean neuron activity of the rest of the population), (iii) SNR of the predicted neuron, (iv) variance across stimuli (computed for mice only, given that there were 32 different stimuli presented in all stimulus recordings in mice; macaque stimulus recordings included repeating the same checkerboard image), and (v) split-half r (Methods). We plotted EV against each of these variables (mouse: ***Figure 6***B, macaque: ***Figure 6***E) and report the correlation coefficient between EV and each variable in the y-axis in ***Figure 6***A (mouse) and ***Figure 6***D (macaque).

In mice, during both stimulus presentation and gray screen presentation, the most correlated property with a neuron's inter-areal predictability was the max r^2 (***Figure 6***A). For the other 4 properties, there was a strong distinction between stimulus presentation (dark bars) and gray screen presentation (light bars): All 4 properties were positively correlated with the neural activity predictability EV fraction during stimulus presentation but they were slightly anticorrelated with their predictability EV fraction during gray screen presentation. Because the split-half correlation calculation averages out the non-stimulus-dependent variability in both halves of the trials, it showed a weaker correlation with EV, which depends on trial-by-trial modulation. The one-vs-rest r^2 metric, which also examines trial-by-trial modulation and does not average split-half trials, yielded a stronger correlation with EV.

When examining the relationship between 1-vs-rest self-consistency and inter-laminar prediction EV in mice, we observed a bimodal distribution of neurons: one group of neurons showed high EV despite having low self-consistency and in the other group EV correlated well with self-consistency (***Figure 6***B third column). The responses of neurons with low self-consistency also showed high EV during gray screen presentation. This bimodality was present in two out of the three mice we tested (MP031 and MP032; self-consistency and EV fraction relationships across all mice can be seen in ***Figure Supplement 5***A). To better understand the responses of neurons with low self-consistency we projected out the "non-visual ongoing neuronal activity" from the neuronal responses (***Stringer et al., 2019a***) (Methods). This non-visual ongoing activity is deemed to be influenced by spontaneous behavior (***Stringer et al., 2019b***). Projecting out this non-visual activity largely led to a unimodal distribution (***Figure 6***C). Removing the non-visual ongoing activity also increased the correlation between self-consistency and inter-laminar predictability (***Figure 6***C). This observation could be because the responses of neurons with low self-consistency can no longer be predicted, or because the responses of those neurons became more reliable and therefore were highly predicted. To distinguish between these two possibilities, we compared both the 1-vs-rest self consistency and the prediction EV before and after removing the non-visual activity. Removing the non-visual ongoing activity increased the self-consistency value across the three mice (***Figure Supplement 5***C; $p < 0.001$, paired permutation test). Interestingly, the inter-laminar EV fraction decreased in MP031 and MP032 mice, yet increased in MP033 (***Figure Supplement 5***D $p < 0.001$, paired permutation test). When examining individual pairwise relationships in a fraction of highly predictable neurons, we found that some of the highly predictable neurons remained predictable after removing the non-visual activity whereas other highly predictable neurons dropped EV fraction dramatically.

In macaques, one of the highest correlated property with inter-areal prediction EV across all conditions was also the max correlation value (***Figure 6***D, E first column). Other neuron properties like SNR, split-half correlation and one-vs-rest correlation were also highly correlated with inter-cortical predictability EV(***Figure 6***D). Unlike the mouse, the split-half correlation was highly correlated with

EV fraction, although the relationship was highly non-linear (***Figure 6***D, middle column). Using the one-vs-rest squared correlation removed some of this non-linearity and further increased the correlation between it and the EV fraction (***Figure 6***D, third column). In addition, there was no bimodal distribution of neurons when relating one-vs-rest correlation and EV fraction.

We conjectured that neurons that have overlapping receptive fields (RFs) should share more information, and therefore their responses would better predict each other than neurons with non-overlapping RFs. In addition, even when all neurons are exposed to the same stimulus (full field symmetrical checkerboard image, gray screen, darkness, etc), neurons with overlapping RFs may be more synaptically connected, resulting in better inter-cortical predictions. To test this hypothesis, we compared inter-cortical predictions in different ensemble of neurons with RFs that differed in their degree of overlap. This hypothesis was only tested in the macaque data because we did not have access to RF estimates in the mouse data. For each V4 site whose responses we predicted, we separated the predictors into two size controlled groups: one where all the V1 predictor sites had <10% RF overlap (sample of one V4 site, ***Figure 6***F, top), and one where all the V1 predictor sites had >80% RF overlap (sample of one V4 site, ***Figure 6***F, bottom). A similar procedure was followed when predicting the activity of V1 neurons from V4 predictor neurons (***Figure 6***H). Inter-areal prediction was higher in the >80% RF overlap condition compared to the <10% RF overlap ensembles in both directions and across all stimulus conditions (***Figure 6***G,I, $n = 110$ total V4 site recordings across all conditiosn, $n = 970$ total V1 site recordings across all conditions). In most cases, predicting the >80% RF overlap ensembles was still lower than ceiling performance (when using all predictors with all types of overlap percentages; ***Figure 6***E,F)

Inter-areal predictability is both stimulus and non-stimulus driven

The results in ***Figure 5*** and ***Figure 6*** pointed to components of the predictable responses that are stimulus driven and other components that are non-stimulus driven. To further examine the non-stimulus driven component, we reasoned that if the shared information between areas were strictly driven by the visual stimulus, then using the activity of a stimulus presentation repeat to one specific image could be used to predict the responses to any other stimulus repeat of the same image. On the other hand, if the shared activity does not have any stimulus response information, then the prediction model would not work when considering responses across repeated presentation of identical stimuli in different trials. To test these two opposing ideas, we compared the inter-areal prediction EV fractions using unshuffled versus shuffled trials. Shuffling was done across repeat trials of the same images (mice: ***Figure 7***A, macaques: ***Figure 7***D). In mice, one stimulus presentation was either a drifting grating or a natural static image. In macaques, one stimulus presentation was either the one checkerboard image, a large thick fast moving bar, or a small thin slow moving bar. In both species and in both directions, inter-areal prediction EV fraction persisted ($p < 0.001$, paired permutation test of shuffled trials prediction vs.shuffled frames prediction), yet the EV fraction decreased after shuffling stimulus repeats compared to before shuffling (***Figure 7***B,E). In mice, neurons showed a bimodal distribution in terms of their response predictability in shuffled and unshuffled trials. For a subset of neurons, the EV fraction was still high in the shuffled condition, albeit their EV was still higher in the unshuffled case (***Figure 7***C; points below but near the diagonal line). For another subset of neurons the EV fraction during shuffled trials was much lower or even near zero. The responses of the latter group had the highest predictability during gray screen activity. In the macaque, there was no bimodal distribution, yet neurons farther away from the diagonal line also had a higher EV fraction during gray screen activity (***Figure 7***F). In addition, we examined whether shuffling repeat presentations of gray screen images (simulating spontaneous activity) would result in any prediction at all. We found a more profund decrease in inter-cortical performance (***Figure Supplement 6***B) with no neurons that remained as predictable during shuffled repeats compared to unshuffled repeats (***Figure Supplement 6***C,D).

Accounting for latency differences improves inter-areal activity predictions in macaque visual area sub-populations

Given the latency differences in neuronal responses between V1 and V4 ***Schmolesky et al., 1998***, we asked whether accounting for this latency could result in better inter-area prediction. To test this hypothesis, we offset the neuronal activity using different lags for each area (***Figure 7***G, H) and recalculated the ridge regression predictions. For each offset level, we calculated the percentage of neurons where the EV fraction peaked at that offset. For the checkerboard image, in the macaque V1→V4 predictions, the biggest percentage of neurons had a peak performance when there was no time offset between areas (***Figure 7***I, left). A substantial proportion of neurons had a peak performance for 25 ms or 50 ms offsets in the negative direction (i.e., V1 activity preceding V4 activity). This distribution of peak EV values was only present during early visual responses (first 275 ms of stimulus onset). In the macaque V4→V1 direction, there was a large proportion of neurons with peak EV when considering 25 ms to 50 ms offsets in the positive direction (i.e., V4 after V1, ***Figure 7***I, right). These differences were apparent in the early part of the visual response, before 250 ms. When offsetting the neuronal responses to gray screen presentations, across all times and areas, the highest percentage of neurons with peak EV was when there was no time offset (***Figure Supplement 6***E,F).

Discussion

Neuronal activity in one brain region or layer within the visual cortex can be used to predict neuronal activity in another nearby and anatomically connected region or layer in single trials (***Figure 2***). In monkeys, predictability was asymmetric: V1 activity better accounted for V4 activity than vice versa (***Figure 3***, ***Figure 7***). This inter-areal prediction persisted across different stimuli (***Figure 4***) but also in the absence of a visual stimulus, during gray-screen and lights-off periods (***Figure 5***). The degree of predictability increased with signal-to-noise ratio, response variance, and the degree of overlap between receptive fields (***Figure 6***).

In line with other studies in mice (***Stringer et al., 2019b***; ***Niell and Stryker, 2008***; ***Andermann et al., 2011***), we observed an approximately bimodal distribution of neuronal responses, with a large subset of neurons that do not show reliable responses to visual stimuli both in L4 and L2/3. Yet, even if these neurons are "non-visual", at least within the set of stimuli and conditions examined here, their activity remains highly predictable. This bimodal distribution dissipates when projecting out potential non-sensory ongoing activity (***Stringer et al., 2019b***, ***2021***). At the population level, neuronal encoding subspaces in mouse visual cortex have been shown to have little overlap between visual sensory and non-sensory (behavioral) information, with only one shared dimension (***Stringer et al., 2019b***). The visually unreliable, yet highly predictable, subset of neurons described here could be the neuronal group driving this orthogonality. As expected, the activity of "visual" neurons can be better predicted during visual presentation and is predicted almost at chance levels during gray screen presentation. In stark contrast, the activity of non-visual neurons can be predicted even better during gray screen presentation than during visual stimulation. There was no such bimodal distribution in the data from monkeys. One possibility is that there may be no (or very few) non-visual neurons in macaque V1 or V4. Indeed the overwhelming majority of neurons in V1 and V4 responded strongly to visual stimulation. Yet, the comparisons between the results in mice and monkeys reported here need to be interpreted with caution because the two datasets differ in terms of recording techniques (electrophysiology versus two-photon imaging), consequently also the temporal resolution (one millisecond versus hundreds of milliseconds), and the type of interaction studied (across areas versus across layers), in addition to any differences between species.

In macaques, sites where the receptive fields (RF) of V1 and V4 overlap can better predict each other compared to other sites showing little RF overlap. This observation could reflect RF-dependent intrinsic connectivity between areas, but also RF-dependent shared inputs from other

areas like the thalamus. In the latter case, those putative shared inputs cannot be strictly dependent on visual inputs given that the effect of RF overlap persists even during gray screen conditions.

Many computational models that aim to explain neuronal activity in visual cortex are based on feedforward signal propagation, with increased receptive field sizes, selectivity, and feature invariance along the visual hierarchy (***Serre et al., 2007b***; ***Kreiman, 2021***; ***Connor et al., 2007***). Consistent with this idea, we described an asymmetry in the degree of predictability, with V1 neurons explaining V4 responses better than the other way around. This observation persisted after controlling for neuronal count and split-half correlation values and was also apparent during the lights-off condition. In contrast, there was no asymmetry when comparing inter-laminar prediction directions in mice. The lack of asymmetry in inter-laminar prediction directions in mice could be due to the slow dynamics in calcium imaging, the lack of a clear inter-areal hierarchy, or differences between species.

The asymmetry in directionality is also observed when implementing temporal delays to inter-areal prediction, consistent with processing delays across areas (***Semedo et al., 2022***; ***Gokcen et al., 2022***; ***Schmolesky et al., 1998***). A substantial proportion of neurons increased their inter-areal predictability when offsetting the times between areas, specifically in the direction that aligns their neuronal activities In contrast to the temporal delays associated with processing visual stimulation, during gray screen presentation, the majority of neurons was best predicted in the absence of any time offsets, suggesting that the internally generated neuronal activity during spontaneous conditions may be largely driven in a non-feedforward manner.

Further evidence supporting the distinction between visually-driven and non-visually-driven interactions comes from the observation that trial repeat shuffling reduced, but did not eliminate, predictability in both mice and monkeys. In mice, when plotting shuffled vs. unshuffled activity, we encountered again a bimodal distribution, where a group of neurons was closer to the diagonal line (their responses were predicted as well during the shuffled compared to the non-shuffled condition), and another group of neurons which were closer to the x-axis (their responses could not be predicted during the shuffled condition). The responses of the latter group were best predicted during gray screen activity, suggesting that they mostly shared non-visual information. The predictive power in mouse V1 from layer 4 to layer 2/3 during spontaneous conditions has been recently shown in (***Papadopouli et al., 2024***), consistent with our findings. The overall area population decrease in predictability after shuffling may be due to the influence of non-visual activity such as movement (***Stringer et al., 2019b***), especially since these non-visual stimulus effects have been shown to occur in the one-second timescale as in our study. In the macaque, context-dependent effects are likely not due to movement, since the monkey maintained fixation during the stimulus task, and visually-evoked activity is not driven by movement (***Talluri et al., 2023***).

The results on the prediction of neuronal responses constitute a lower bound. First, we focus on linear predictability but other (non-linear) models could better capture neuronal activity. Second, and critically, the experimental data provide only a fraction of the inputs to a given neuron–excluding most (in macaque dataset) if not all (in mouse dataset) inhibitory inputs that are crucial for the organization of circuit and microcircuits in visual cortex (***Jiang et al., 2015***; ***Shen et al., 2020***; ***Ibrahim et al., 2020***; ***Schuman et al., 2021***). Third, biophysical realistic models of the transformation between inputs and outputs of a given neuron should include their dendritic locations and specific synaptic potentials (***Park et al., 2019***; ***Petousakis et al., 2023***).

We introduce a unifying method to evaluate inter-areal interactions in different types of neuronal recordings, timescales, and species. These interactions can be assessed in single trials, separating visually-driven and non-visual contributions, and accounting for the directionality and dynamics of neuronal responses. These efforts constitute an initial step towards systematically building computational models that can account for the transformations from sensory inputs to their encoding in the cortex.

Methods

Datasets. We used the mouse dataset from (***Stringer et al., 2019a***) containing calcium-imaging activity measurements from 43,630 neurons in layer 4 (L4) and 12,060 neurons in layers 2/3 (L2/3) in V1 of 4 mice during 32 different randomly interleaved presentations of either drifting gratings or gray-scale natural images (each one repeated more than 90 times), along with spontaneous activity during 30 minutes of exposure to a gray/black screen (***Figure 1***A, data acquisition details in (***Stringer et al., 2019a***,b)). Calcium imaging activity was recorded during stimulus presentations at a scan rate of 2.5 Hz or 3 Hz (each frame was acquired every 400 ms or 333 ms). The computed stimulus responses per stimulus presentation were averaged based on two frames immediately post stimulus onset. Cortical layers were determined using the 10-12 planar z-positions retrieved during the multi-plane calcium activity acquisition. For stimulus-response and spontaneous recordings, neuronal activity of each neuron was z-scored using its 30-minute gray screen spontaneous activity (mean gray-screen activity subtracted and divided by gray-screen activity standard deviation).

We used the macaque monkey dataset from (***Chen et al., 2022***). This dataset consists of envelope multiunit activity (MUAe) from 1,024 recording sites in one monkey in response to either multiple-day recordings of more than 60 repetitions of a full-size checkerboard image, moving small-thin or large-thick bars in 4 directions, gray screen presentations, or more than 30 minutes of baseline activity where the monkey was in a room with the lights off (***Figure 1***B). Neuronal activity was averaged over 25 ms non-overlapping bins. Activity duration was 300 ms, 400 ms, and 1 s for gray screen, checkerboard, and moving bar presentations, respectively. For the recordings during visual stimulation, the neuronal activity was normalized by subtracting the mean activity during the gray screen presentations separately for each site.

Visual responsiveness. A neuron or site was defined to be visually responsive if its signal-to-noise ratio (SNR) was 2 or higher and its split-half correlation value was 0.8 or higher. Due to the high number of repetitions of visual stimuli, the split-half correlation was skewed toward high values, which is why we used a higher split-half correlation threshold than commonly used in other studies. In mice, the SNR for each neuron was calculated as:

$$SNR_{mouse} = \frac{< r_{stim} > - < r_{spont} >}{std(r_{spont})} \tag{1}$$

where <> denotes mean, std denotes the standard deviation, r_{stim} is the average activity in response to stimuli, and r_{spont} indicates the average activity over the 30-minute gray screen presentation activity.

In monkeys, we followed the definition in (***Chen et al., 2022***) and calculated the SNR for each site as:

$$SNR_{monkey} = \frac{max(< r_{stim} >) - < r_{spont} >}{std(r_{spont})} \tag{2}$$

using the peak activity during the checkerboard presentation for the signal, and the average gray screen neuronal activity as background (denoted as r_{spont}).

In mice, the split-half consistency was calculated by correlating the average activity for the 32 stimuli in a randomly chosen half of the trials, with the average activity in the other half of the trials, followed by Spearman-Brown correction (used to correct for the division of trials by half). In monkeys, during checkerboard presentations, the split-half consistency was calculated by correlating the average activity of the 16 timepoints (0–400 ms; 25 ms bins) of checkerboard presentation of 25 random trial repetitions with the average activity of another non-overlapping 25 random repetitions, followed by Spearman-Brown correction. During moving bar presentations, the 40 timepoints (0–1s; 25 ms bins) during 25 random trial repetitions were first concatenated across the 4 directions (total of 160 timepoints), and then correlated to the concatenated averaged activity of another nonoverlapping 25 random trial repetitions, followed by Spearman-Brown correction. For all split-half consistency calculations, we randomly sampled trials 20 times.

Inter-areal regression. Let ${}_{A}r_{i,t}$ be the activity of neuron or site i in area A at timepoint t, where A can be L4 or L2/3 in the mouse data and V1 or V4 in the monkey data. Neuronal activity from one area (e.g., mouse V1 L4 or macaque V1) was used to predict activity in the other area (e.g., mouse V1 L2/3 or macaque V4) using ridge regression (***Figure 1***E). The activity of each neuron i in area A2 was predicted from n_{A1} neurons in area A1 as follows:

$$ {}_{A2}\hat{r}_{i,t} = \sum_{j=1}^{n_{A1}} w_{i,j\ A1} r_{j,t} + b_i \tag{3} $$

During fitting, we minimized the residual sum of squares (RSS), defined as:

$$ RSS_i(\mathbf{w}, b_i) = \sum_{t=1}^{n_T} ({}_{A2}\hat{r}_{i,t} - {}_{A2}r_{i,t})^2 + \alpha \sum_{j=1}^{n_{A_1}} w_j^2 \tag{4} $$

where **w** is the weight vector for predicting the activity of neuron i, n_T is the number of images/time points and α controls the regularization strength (α was tuned for each dataset with an independent sample and ranged from 10^3 to 10^5). Predictability for each neuron was evaluated using 10-fold cross-validation across trials and quantified as squared Pearson's r, referred to as explained variance fraction (EV fraction) throughout.

To remove temporal auto-correlation that would inflate the apparent prediction despite cross-validation, we removed training timepoints near the test timepoints closer than the decay window of the activity auto-correlation (mouse: 5 s; macaque: 1.5 s). The auto-correlation decay window was determined using time-series forecasting Ridge Regression (using r_t to predict r_{t+d}, where d represents a delay). The delay was increased until the EV fraction approached chance.

Prediction directionality. We compared predictability across layers in different directions (in mice: L4→L2/3 vs. L2/3→L4) and also predictability across areas in different directions (in macaques: V1→V4 vs. V4→V1) (***Figure 3***). To ensure that results were not dependent on the number of neurons/sites, we randomly subsampled the number of neurons/sites of the area containing the larger number of neurons/sites (L2/3 for mouse; V1 for macaques) to match the number of predictors in both directions (10 permutations, neuron count details in ***Table 1*** and ***Table 2***). To account for potential changes in intrinsic predictability, we ensured that the neurons from both areas were matched in terms of the distribution of split-half correlation values so that the difference between individual area neurons/sites was less than 0.002. To assess the intrinsic predictability of neurons/sites in each region, the areas were used to predict themselves, where one neuron/site in the area was predicted by the remaining neurons in the same area. This "intra-areal" prediction was used to normalize EV fraction to compare directionality of prediction.

Stimulus types and spontaneous activity comparison. We compared predictability for different stimulus conditions ***Figure 4***, ***Figure 5***). To compare inter-areal prediction across stimulus types and between the presence or absence of stimuli , the number of predictors (neurons or sites) and timepoints was sub-sampled to be the same across all datasets. In the macaque, the time spent recording the lights-off condition was much greater than during stimulus or gray screen presentations. To account for the difference in time duration and therefore training size, we subsampled time periods to be the same across all stimulus, gray-screen, and lights-off, lights-off eyes open, and lights-off eyes closed conditions.

Neuron properties. We compared different neuronal properties with predictability measurements (***Figure 6***). The SNR and split-half correlation has been defined above. The absolute max pairwise correlation value of each neuron/site i in one area with all neurons in the other area was calculated as

$$ {}_{A2}maxcorr_i = max_j \left(|corr({}_{A2}r_i, {}_{A1}r_j)| \right) \tag{5} $$

where ${}_{A2}r_i$ represents the activity of neuron/site i in area A2, which are correlated with the activity of every neuron j in area A1 (denoted as ${}_{A1}r_j$).

The one-vs-rest correlation was calculated as follows. In the mouse data, we correlated the activity for the 32 stimuli during 1 trial repetition with the averaged activity of the remaining trial

repetitions. In monkey, during checkerboard presentations, the one-vs-rest correlation was calculated by correlating the activity of the 16 timepoints (0–400 ms; 25 ms bins) during 1 trial repetition with the averaged activity of the remaining trial repetitions. For moving bar presentations, the 40 timepoints (0–1s; 25 ms bins) during 1 trial repetition were first concatenated across the 4 directions (total of 160 timepoints), and then correlated to the concatenated, trial-averaged activity of the remaining trial repetitions. For all one-vs-rest correlation calculations, we held out each trial in turn and averaged across the samples.

Receptive field overlap comparisons. In macaque, receptive field (RF) ellipses were calculated using center and edge locations in the dataset (***Figure 6***E, F). To calculate the percentage of RF overlap between the neuronal sites to be predicted and the predictor, the intersection area between ellipses was retrieved using the Shapely python package, and divided by the area of the predicted site. Sites that had predictors that overlap both more than 80% and less than 10% were selected to compare inter-areal predictions. To control for predictor size, 14 random predictor sites from all the sites in each overlap type were subsampled (10 random samples without replacement).

Trial repeat shuffling and time offset predictions. For the shuffled-trial experiments, we shuffled the predictor activity across repeat trials showing the same stimulus (***Figure 7***). (Thus, the stimulus order remained the same.) For the mouse time-offset analysis, the activity of predictor neurons was time-shifted in the positive or negative direction, with 1 bin corresponding to 1 stimulus presentation (800–900ms). For the monkey dataset, the predictor activity (400 or 1000 ms per presentation, 16–50 bins of 25 ms each) was offset across time bins. We used sub-windows of 200 ms to avoid window-length differences that would otherwise be introduced if we shifted the entire trial response.

Data and code availability. All the computational models and data analysis code developed in this work is publicly available at this link: https://github.com/4sdch/inter-area-neural-prediction.git. All the data are publicly available for mouse: https://figshare.com/articles/dataset/Recordings_of_ten_thousand_neurons_in_visual_cortex_in_response_to_2_800_natural_images/6845348?file=12462734 (***Stringer et al., 2019a***) and for macaques: https://gin.g-node.org/NIN/V1_V4_1024_electrode_resting_state_data (***Chen et al., 2022***).

Figures and Tables 503

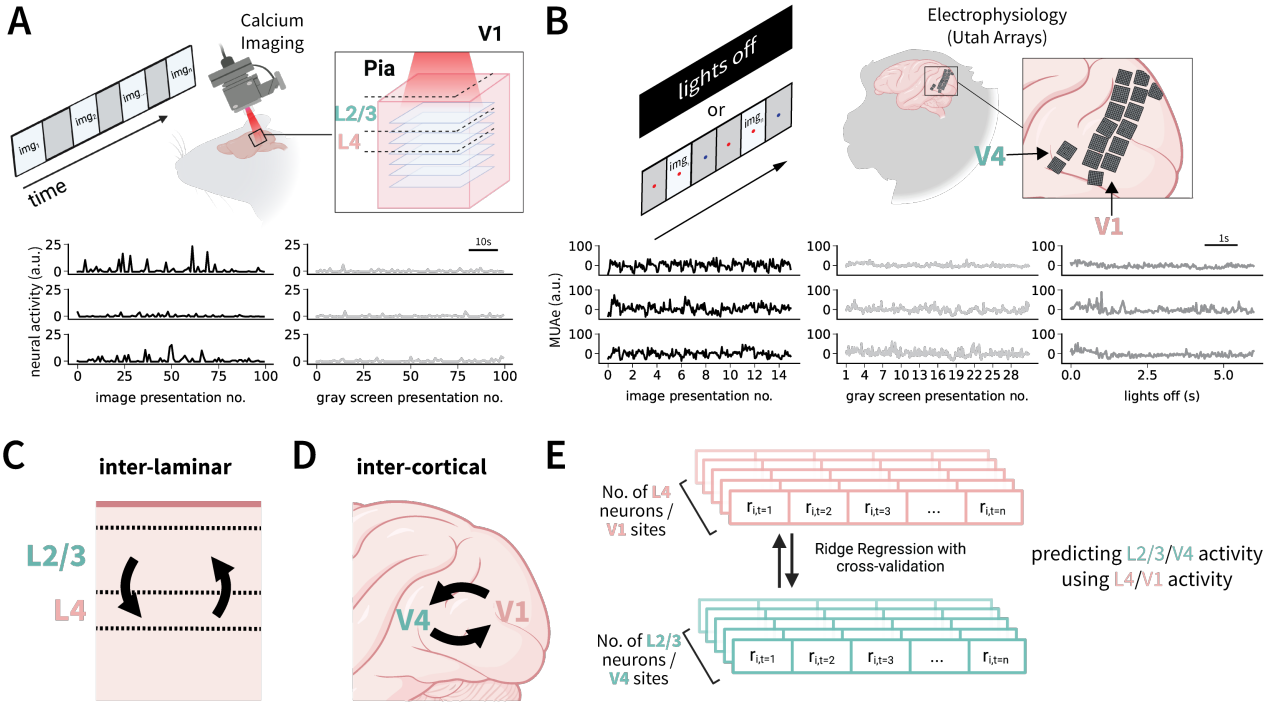

Figure 1. Predicting trial-to-trial and timepoint-to-timepoint neuronal activity between areas. A. Top: Experimental set-up to record two-photon Calcium imaging activity data from layers 2/3 (L2/3) and layer 4 (L4) in rodent V1 upon presentation of gratings, natural stimuli or gray screen images (represented as img_n) (***Stringer et al., 2019a***). Deconvolved calcium imaging traces were z-scored using baseline activity during 30 minutes of gray screen presentation before/after image presentation (***Table 1***). Bottom: Sample z-scored neuronal activity from 3 different neurons in response to 100 presentations of drifting gratings (left) or gray screen presentations (right). Each activity value corresponds to one image presentation, and calculated as the average of two calcium imaging video frames (666 ms or 800 ms, see details in Methods). **B.** Top: Experimental set-up for the neuronal activity data from macaque V1 and V4 (***Chen et al., 2022***). Electrophysiological activity was simultaneously recorded across 1,024 channels from 16 Utah arrays (***Table 2***). Bottom: Envelope multiunit spiking activity (MUAe) from 3 different sites in response to multiple presentations of a repeated 400 ms full-field checkerboard image (left, baseline mean-subtracted), 200 ms gray screen (middle), or during a lights-off condition (30 minutes total; right). Each value corresponds to aggregated MUAe activity in a 25 ms bin. **C.** Overview of inter-laminar relationships examined in mouse V1. “lower-level" layer 4 (L4) neuronal activity is used to predict “higher level" layer 2/3 (L2/3) activity (blue arrow) and vice versa (red arrow). **D.** Overview of inter-cortical relationships examined in macaque, where lower-level V1 is used to predict higher-level V4 (blue arrow) and vice versa (red arrow). **E.** Illustration of linear ridge regression method used for inter-areal prediction. Neuronal activity in response to presentation number i (labeled “r_i”) at time t from one area (e.g., mouse V1 L2/3 or macaque V1) was used to predict activity in the other area (e.g., mouse V1 L4 or macaque V4) (***Semedo et al., 2019***). Predictability was evaluated using 10-fold cross-validation across presentation trials in mouse, and across 25 ms timepoints in macaque (Methods).

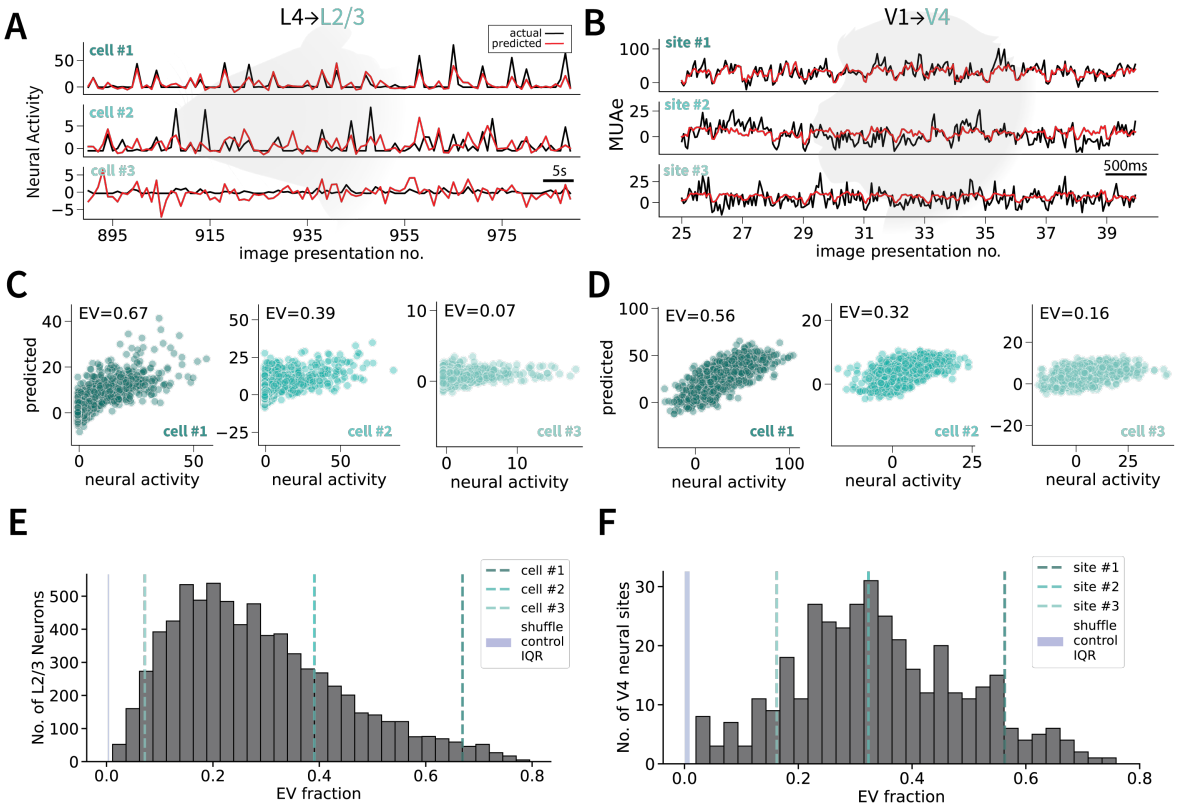

Figure 2. Lower level activity can predict higher level activity in both rodent and primate brains. **A.** Example neuronal activity (z-scored, black) in response to stimulus presentations (drifting gratings) in mouse V1 L2/3 along with regression-model predictions (red) for a typical cell (2, middle), cell in the top 10% percentile of predictability (1, top), and bottom 10% percentile (3, bottom). **B.** Same as **A** for macaque MUAe activity in response to a full-field checkerboard image in three V4 neuronal sites. **C.** Predicted neuronal activity versus actual neuronal activity in response to stimuli for the mouse L2/3 cells 1, 2, and 3 shown in **A**. Each point represents 800 ms corresponding to a stimulus presentation. r values (top left) indicate the correlation coefficient. **D.** Same as **C** for macaque V4 neuronal sites 1, 2, and 3. Each point represents one 25ms timepoint during the 400 ms presentation. **E**. Distribution of EV fraction in L4→L2/3 regressions of neural activity in response to visual stimuli in cells that were deemed visually responsive in 4 mice and 7 recording days ($n = 7,265$ neurons, Methods). Performance using 10-fold cross-validation across trials was quantified as squared Pearson's r, referred to as explained variance (EV) fraction. The three vertical lines show the 3 examples in part **A, C**. The blue solid shaded rectangle (here and throughout) represents the interquartile range (IQR) shuffle control performance, where the activity timepoints of one area were randomly shuffled. **F.** Distribution of EV fraction in V1→V4 regressions of neural activity in response to vistual stimuli in sites deemed visually responsive (One macaque, 5 recording days, 68–82 V4 sites recorded per day; $n = 376$ total site recordings).

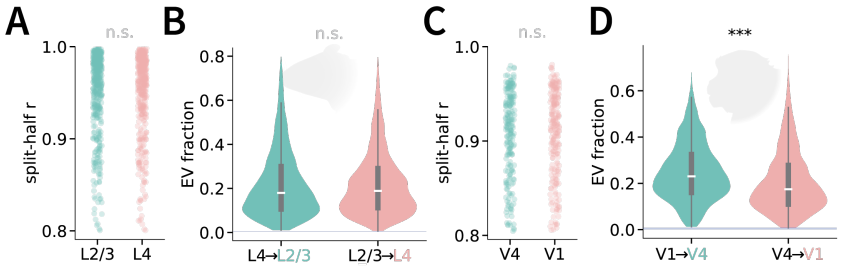

Figure 3. Asymmetry in inter-cortical predictability in macaque but not inter-laminar predictability in mouse. **A.** Split-half reliability (Methods) for the $n = 298$ neurons (per area) in mouse MP033 drifting gratings presentation recording of V1 L2/3 (green) and L4 (coral) used to perform directionality comparisons. Neurons were randomly sub-sampled to match the numbers and self-reliability in the two distributions. Here and throughout, asterisks indicate statistically significant differences using a hierarchical independent permutation test (10,000 permutations): * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; "n.s." indicates $p > 0.05$. **B.** Violin plots describing the distribution of EV fraction for L4→L2/3 (green) and L2/3→L4 (coral) predictions across all 7 stimulus recordings ($n = 1,113$ neurons per area). Violin plots (here and throughout) represent the distribution of neuron/site values, with width representing density and inner boxplot representing the interquartile range. Whiskers of innerbox represent range of the data. **C.** Split-half reliability for the $n = 74$ sites (per area) in macaque checkerboard recording (date=090817) of V4 (green) and V1 (coral) used to perform directionality comparisons. **D.** Violin plots describing the distribution of EV fraction for V1→V4 (green) and V4→V1 (coral) across all 5 stimulus recordings ($n = 786$ sites recordings per area).

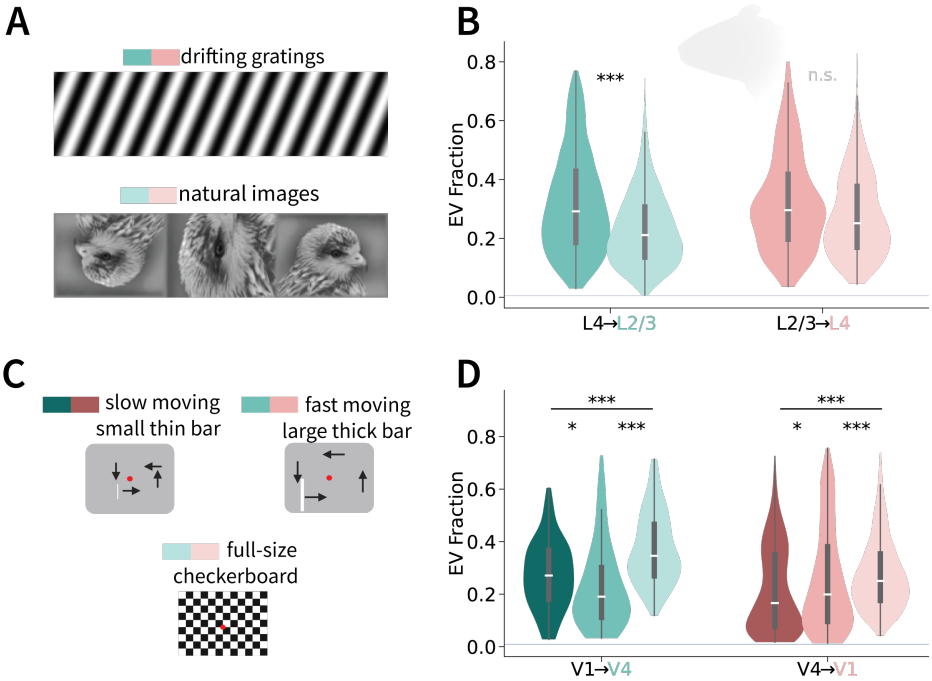

Figure 4. Stimulus type influences neuronal predictability. A. Illustration of the two types of stimuli (drifting gratings and static natural images) presented to the mouse during calcium imaging. **B.** Across-layer predictability in mouse V1 for each stimulus type (dark: drifting gratings, light: natural images) and prediction direction. **C.** Illustration of the three types of stimuli presented to the monkeys (***Chen et al., 2022***). The slow-moving small thin bar moved near the fixation point for 1 s in each of the four directions, while the fast moving large thick bar moved towards the edges of the screen for 1 s in each of the four directions. The full-field checkerboard image was presented repeatedly (400 ms each presentation). **D.** Across-area predictability for each stimulus type (dark: slow bars, medium: fast bars, light: checkerboard) and direction.

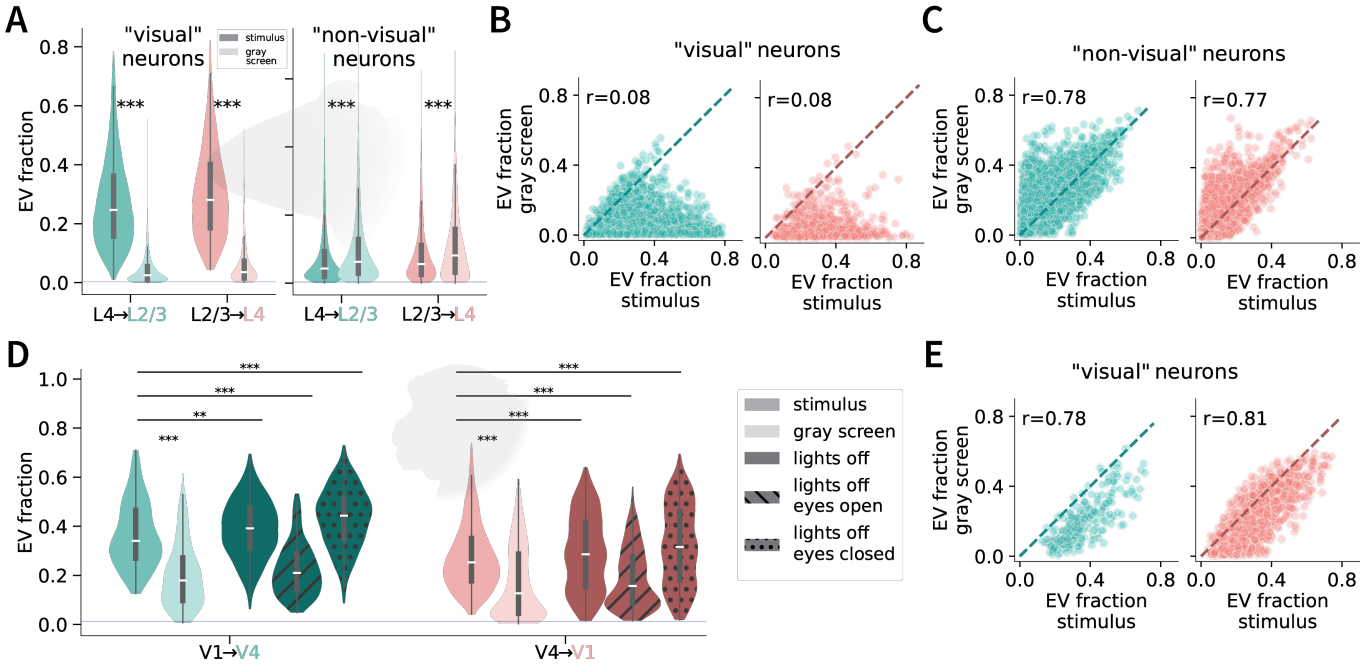

Figure 5. Spontaneous activity can also be predicted. A. EV fraction of neuronal activity in response stimulus presentation (dark violins) or gray screen presentation (light violins) for neurons deemed visually (left) or non-visually (right) responsive (Methods). **B.** Correlation between EV in responses to gray screen (y-axis) versus stimulus presentation (x-axis) in mouse V1 visually responsive neurons (L4→L2/3:left, green; L2/3→L4: right, coral). Diagonal line represents the line of equality (y=x). r is the Pearson's r coefficient. **C.** Same as B but for non-visually responsive neurons. **D.** EV during stimulus presentations (checkerboard image, green), gray screen presentations (light green) or during lights off (dark green). The lights-off condition is further separated into periods when the eyes were open or closed. All lights-off conditions were sub-sampled (10 permutations) to contain similar training lengths as the stimulus and gray screen presentation recordings. **E.** Correlation between EV in responses to gray screen (y-axis) versus stimulus presentation (x-axis) in macaque visually responsive neurons (V1→V4:left, green; V4→V1: right, coral). Diagonal line represents the line of equality (y=x). r is the Pearson's r coefficient. All recorded sites were pulled from the 3 recording days of checkerboard presentations.

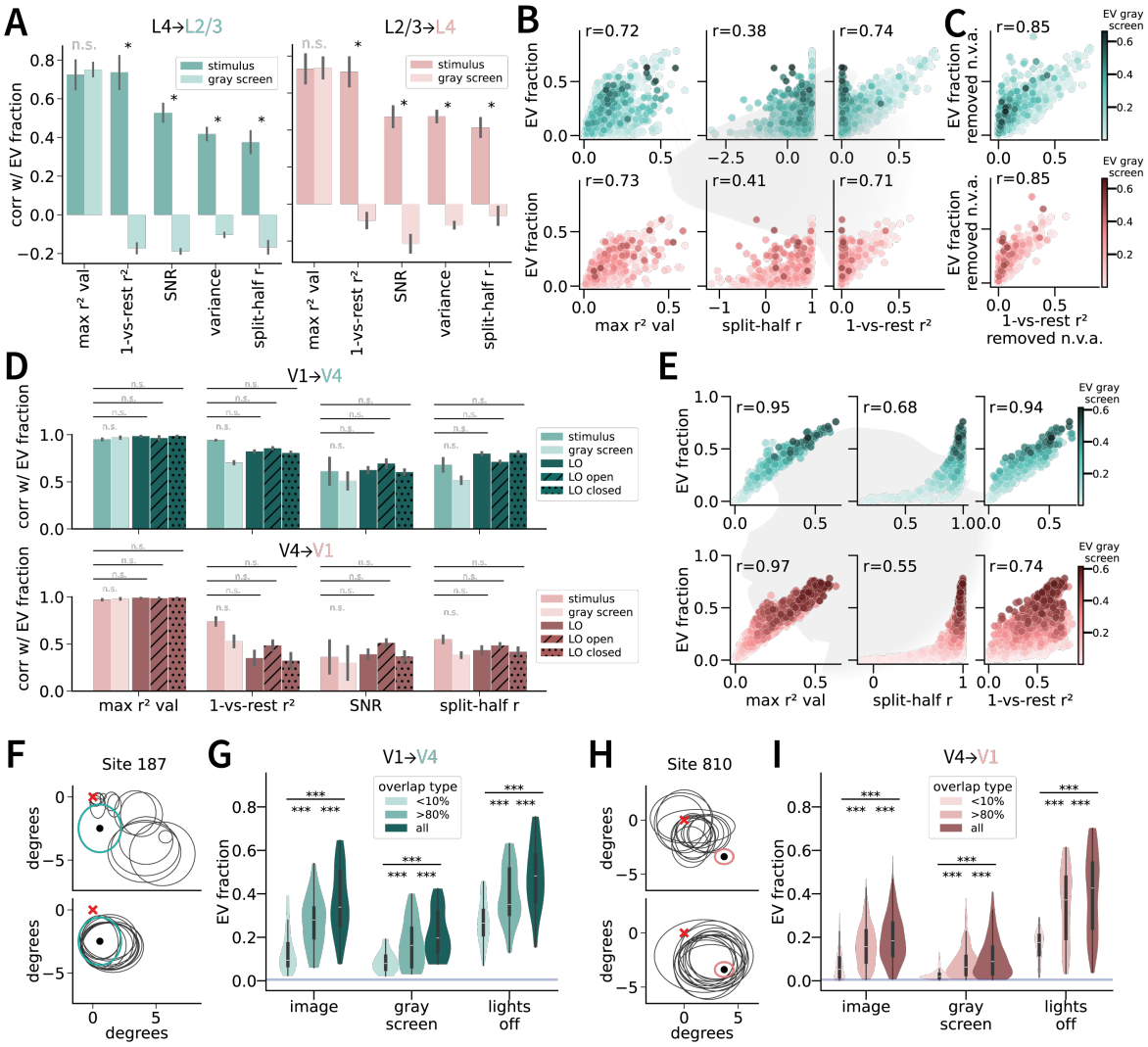

Figure 6. neuronal predictability depends on SNR, stimulus response variance, and receptive field overlap. A. Correlation between different neuronal properties with the predictability of L2/3 (green) and L4 (coral) neuronal responses during the presence (dark color) or absence (light color) of visual stimulus. Neuronal properties measured in mouse V1 include the correlation value of the most correlated pair to each cell (max correlation value, squared), a modified metric of self-consistency (one-vs-rest correlation, squared), SNR, variance in the neuronal activity in response to different stimuli, and the traditional metric of self-consistency (split-half correlation r) (Methods). **B.** Relationship between three neuronal properties and their predictability in a randomly chosen sub-sample of neurons ($n = 4,000$) for mouse L2/3 (green) and L4 (coral) neuronal responses from both drifting gratings and natural images conditions (combined). Hue represents the degree of predictability for the same neurons during the 30 minutes of gray screen presentation (see color map on bottom right).**C.** 1-vs-rest square correlation relationship with predictability after projecting out dimensions of "non-visual" activity (using gray screen activity (***Stringer et al., 2019a***). **D.** Correlation between different neuronal properties with the predictability of monkey V4 (green) and V1 (coral) neuronal site recordings during the full-field checkerboard presentation (dark color), gray screen presentation (light color), and lights-off condition (darkest color; solid, hatch lines, and hatch dots). Neuronal properties measured in the macaque visual cortex include the max correlation squared value, one-vs-rest squared correlation, SNR, and split-half correlation r. **E.** Same as **B** for the macaque V1 and V4 neuronal sites. **F.** Top: Receptive fields of one sample V4 neuronal site (green circle, array 2 electrode 187) and 14 randomly selected V1 neuronal sites as predictors (black circles), constrained on sites that share less than 10% receptive field overlap with the V4 site. Bottom: Receptive fields of the same neuronal site 187 and 14 randomly selected V1 neuronal sites used as predictors, constrained on sites that share at least 80% receptive field overlap with the V4 site. **G.** Differences in predictability of V4 neural activity ($n = 110$ site recordings) in terms of 14 V1 predictor sites with less than 10% RF overlap (light green), 14 predictor sites with at least 80% RF overlap (green), and all predictors (dark green). Predictions were computed for recordings in response to the stimulus presentation (sliding bars and full-field checkerboard images), gray screen presentation, and lights off. **H.** Bottom and top left: Same as **F** but for macaque sample V1 site 810. **I.** Same as **D**, but for V1 ($n = 970$ site recordings).

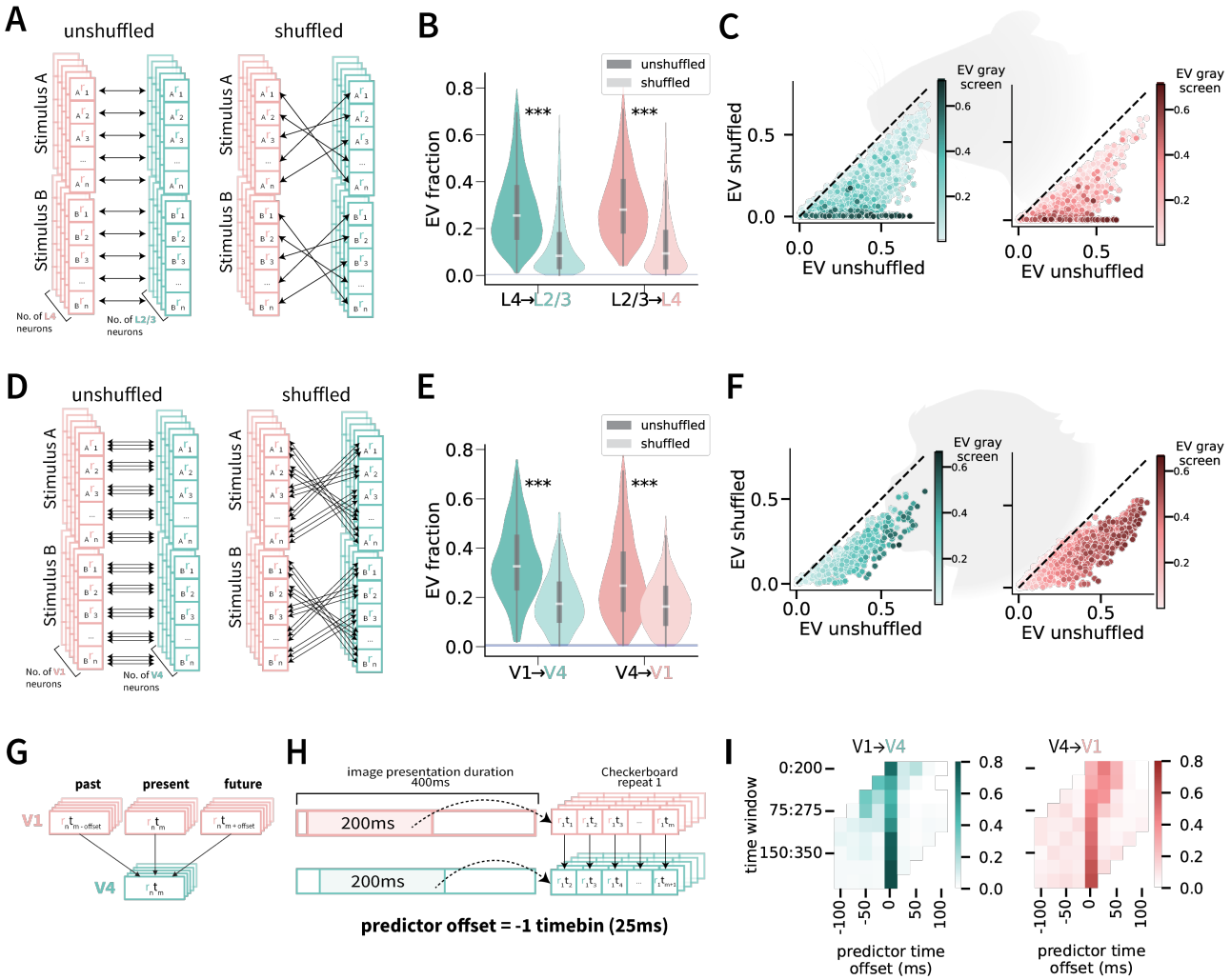

Figure 7. Predicting neuronal activity across time reveals shared stimulus- and non-stimulus driven information in both mouse and macaque visual cortex, along with latency and non-latency specific correlations in macaque V1/V4. A. Shuffled-trial-repeat experiment set-up for comparing unshuffled vs. shuffled prediction in mouse V1 L2/3 and L4. Neuronal activity in response to stimulus repeats was shuffled within their respective image. **B.** EV fraction for unshuffled (dark) and shuffled (light) trials in the L4 → L2/3 (green) and L2/3 → L4 (coral) directions. **C.** Relationship between shuffled (y-axis) and unshuffled (x-axis) trial repeat EVs in the mouse L4 → L2/3 (left, green) and L2/3 → L4 (right, coral) directions. Hue represents EV fraction during gray screen activity. **D.** Shuffled-trial-repeat experiment set-up comparing unshuffled vs. shuffled prediction in macaque data. Neuronal activity (including all timepoints) in response to stimulus repeats were shuffled within their respective image. Since checkerboard presentation was only one stimulus, visualization of experiment only applies to the "Stimulus A" portion. **E.** Same as **B** for macaque V1 → V4 (green) and V4 → V1 (coral). **F.** Same as **C** for macaque V1 → V4 (green) and V4 → V1 (coral). **G.** Illustration of time offsets applied to macaque neuronal activity for inter-areal predictions. Instead of neuronal activity prediction between areas being done on simultaneous activity (middle coral and bottom green box), the V4 neuronal activity (green) at time t_m was predicted using V1 neuronal activity (coral) at time $t_{m\pm offset}$, were *offset* represents 1–8 timebins (25 ms per timebin) before (if negative; left coral box) or after (if positive; right coral box) time t_m. Time offset experiment was done in both prediction directions (V1 → V4 and V4 → V1). **H.** Experimental set-up example for predicting neuronal activity in V4 using V1 neuronal activity from 25 ms prior to V4 activity. Neural activity is in response to a repeated checkerboard image. A 200 ms section of the cortical area was used to represent the image presentation response, and was offset -1 timebin (25 ms) to predict a 200 ms target cortical area. During the prediction experiments, the 200 ms window was slid across the entire duration of the stimulus **I.** Time offset prediction results across both V1→V4 (left,green) and V4→V1 (right, coral) prediction directions. Each square corresponds to the fraction of neuronal sites whose neural activity were best predicted during that offset period and time window.

Table 1. Mouse neuron counts used for inter-layer prediction and analyses. A total of 7 recordings were used to perform prediction experiments. Each row corresponds to a recording day, containing the dataset recording type (Mouse Dataset), total number of neurons, and visually responsive neurons (see Methods). Fourth column: In the directionality prediction experiments, the area containing more neurons (L2/3) was further subsampled to match the number of L4 neurons. The dataset recording type names contain either "ori32" or "natimg32", in addition to the mouse name (MP0-). "natimg32" represents dataset of the 32 natural image presentation. "ori32" represents dataset of the 32 drifting gratings.

	Layer 2/3			Layer 4	
Mouse Dataset	Total	visually responsive	Subsampled (directionality)	Total	visually responsive
natimg32 MP031	6615	1248	219	2367	219
natimg32 MP032	7980	1549	96	1441	96
natimg32 MP033	6646	1467	164	2010	164
ori32 MP027	6264	1029	211	2346	211
ori32 MP031	5423	455	78	1382	78
ori32 MP032	5420	274	47	923	47
ori32 MP033	5277	1243	298	1588	298
Total	43625	7265	1113	12057	1113

Table 2. Monkey site counts used for inter-cortical prediction and analyses. Dates 090817,100817,250717 correspond to neuronal activity in response to checkerboard presentations, gray screen presentations, and lights off condition. Date 260617 corresponds to small thin moving bars presentation. Date 280617 corresponds to large thick moving bars presentation. Fourth column: In the directionality prediction experiments, the area containing more sites (V1) was further subsampled to match the number of V4 sites.

	V1			V4	
Date	Total	visually responsive	Subsampled (directionality)	Total	visually responsive
090817	627	553	74	96	74
100817	688	589	81	112	81
250717	645	537	71	86	71
260617	645	593	82	86	82
280617	645	518	68	86	68
Total	3250	2829	376	469	376

Acknowledgments

The authors would like to thank Elisa Pavarino, Leonardo Polina, Sara Djambazovksa, Carlos Ponce, Jan Drugowitsch, and Wei-Chung Allen Lee for providing comments on the manuscript.

References

Andermann ML, Kerlin AM, Roumis DK, Glickfeld LL, Reid RC. Functional Specialization of Mouse Higher Visual Cortical Areas. Neuron. 2011 Dec; 72(6):1025–1039. https://www.cell.com/neuron/abstract/S0896-6273(11)01012-9, doi: 10.1016/j.neuron.2011.11.013, publisher: Elsevier.

Avitan L, Stringer C. Not so spontaneous: Multi-dimensional representations of behaviors and context in sensory areas. Neuron. 2022 Oct; 110(19):3064–3075. https://www.sciencedirect.com/science/article/pii/S0896627322005888, doi: 10.1016/j.neuron.2022.06.019.

Chacron MJ, Longtin A, Maler L. The effects of spontaneous activity, background noise, and the stimulus ensemble on information transfer in neurons. Network (Bristol, England). 2003 Nov; 14(4):803–824.

Chen X, Morales-Gregorio A, Sprenger J, Kleinjohann A, Sridhar S, van Albada SJ, Grün S, Roelfsema PR. 1024-channel electrophysiological recordings in macaque V1 and V4 during resting state. Scientific Data. 2022 Mar; 9(1):77. https://www.nature.com/articles/s41597-022-01180-1, doi: 10.1038/s41597-022-01180-1, number: 1 Publisher: Nature Publishing Group.

Connor CE, Brincat SL, Pasupathy A. Transformation of shape information in the ventral pathway. Current Opinion in Neurobiology. 2007 Apr; 17(2):140–147. doi: 10.1016/j.conb.2007.03.002.

Consortium M, Bae JA, Baptiste M, Bodor AL, Brittain D, Buchanan J, Bumbarger DJ, Castro MA, Celii B, Cobos E, Collman F, Costa NMd, Dorkenwald S, Elabbady L, Fahey PG, Fliss T, Froudarakis E, Gager J, Gamlin C, Halageri A, et al. Functional connectomics spanning multiple areas of mouse visual cortex. bioRxiv; 2021.

Dadarlat MC, Stryker MP. Locomotion Enhances Neural Encoding of Visual Stimuli in Mouse V1. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience. 2017 Apr; 37(14):3764–3775. doi: 10.1523/JNEUROSCI.2728-16.2017.

Douglas RJ, Martin KAC. Neuronal circuits of the neocortex. Annual Review of Neuroscience. 2004; 27:419–451. doi: 10.1146/annurev.neuro.27.070203.144152.

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex (New York, NY: 1991). 1991; 1(1):1–47. doi: 10.1093/cercor/1.1.1-a.

Gazzaley A, Rissman J, Cooney J, Rutman A, Seibert T, Clapp W, D'Esposito M. Functional Interactions between Prefrontal and Visual Association Cortex Contribute to Top-Down Modulation of Visual Processing. Cerebral Cortex. 2007 Sep; 17(suppl_1):i125–i135. https://doi.org/10.1093/cercor/bhm113, doi: 10.1093/cercor/bhm113.

Gilbert CD, Li W. Top-down influences on visual processing. Nature Reviews Neuroscience. 2013 May; 14(5):350–363. https://www.nature.com/articles/nrn3476, doi: 10.1038/nrn3476, number: 5 Publisher: Nature Publishing Group.

Gokcen E, Jasper AI, Semedo JD, Zandvakili A, Kohn A, Machens CK, Yu BM. Disentangling the flow of signals between populations of neurons. Nature Computational Science. 2022 Aug; 2(8):512–525. https://www.nature.com/articles/s43588-022-00282-5, doi: 10.1038/s43588-022-00282-5, number: 8 Publisher: Nature Publishing Group.

Hohaia W, Saurels BW, Johnston A, Yarrow K, Arnold DH. Occipital alpha-band brain waves when the eyes are closed are shaped by ongoing visual processes. Scientific Reports. 2022 Jan; 12:1194. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8786963/, doi: 10.1038/s41598-022-05289-6.

Hsu A, Borst A, Theunissen FE. Quantifying variability in neural responses and its application for the validation of model predictions. Network: Computation in Neural Systems. 2004 Jan; 15(2):91–109. https://doi.org/10.1088/0954-898X_15_2_002, doi: 10.1088/0954-898X_15_2_002, publisher: Taylor & Francis _eprint: https://doi.org/10.1088/0954-898X_15_2_002.

Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of Physiology. 1962; 160(1):106–154. https://onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1962.sp006837, doi: 10.1113/jphysiol.1962.sp006837, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1962.sp006837.

Ibrahim LA, Schuman B, Bandler R, Rudy B, Fishell G. Mining the jewels of the cortex's crowning mystery. Current Opinion in Neurobiology. 2020 Aug; 63:154–161. https://linkinghub.elsevier.com/retrieve/pii/S0959438820300842, doi: 10.1016/j.conb.2020.04.005.

Jiang X, Shen S, Cadwell CR, Berens P, Sinz F, Ecker AS, Patel S, Tolias AS. Principles of connectivity among morphologically defined cell types in adult neocortex. Science. 2015 Nov; 350(6264):aac9462. https://www.science.org/doi/10.1126/science.aac9462, doi: 10.1126/science.aac9462, publisher: American Association for the Advancement of Science.

Kreiman G. Biological and Computer Vision. Cambridge: Cambridge University Press; 2021. https://www.cambridge.org/core/books/biological-and-computer-vision/BB7E68A69AFE7A322F68F3C4A297F3CF, doi: 10.1017/9781108649995.

Lee WCA, Bonin V, Reed M, Graham BJ, Hood G, Glattfelder K, Reid RC. Anatomy and function of an excitatory network in the visual cortex. Nature. 2016 Apr; 532(7599):370–374. https://www.nature.com/articles/nature17192, doi: 10.1038/nature17192, number: 7599 Publisher: Nature Publishing Group.

Markov NT, Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C, Magrou L, Vezoli J, Misery P, Falchier A, Quilodran R, Gariel MA, Sallet J, Gamanut R, Huissoud C, Clavagnier S, Giroud P, Sappey-Marinier D, Barone P, Dehay C, Toroczkai Z, Knoblauch K, et al. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cerebral Cortex (New York, NY: 1991). 2014 Jan; 24(1):17–36. doi: 10.1093/cercor/bhs270.

Morales-Gregorio A, Kurth AC, Ito J, Kleinjohann A, Barthélemy FV, Brochier T, Grün S, van Albada SJ. Neural manifolds in V1 change with top-down signals from V4 targeting the foveal region. Cell Reports. 2024 Jul; 43(7):114371. https://www.sciencedirect.com/science/article/pii/S2211124724006995, doi: 10.1016/j.celrep.2024.114371.

Niell CM, Stryker MP. Highly Selective Receptive Fields in Mouse Visual Cortex. The Journal of Neuroscience. 2008 Jul; 28(30):7520. https://pmc.ncbi.nlm.nih.gov/articles/PMC3040721/, doi: 10.1523/JNEUROSCI.0623-08.2008.

Niell CM, Stryker MP. Modulation of visual responses by behavioral state in mouse visual cortex. Neuron. 2010 Feb; 65(4):472–479. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184003/, doi: 10.1016/j.neuron.2010.01.033.

Okazaki Y, Abrahamyan A, Stevens CJ, Ioannides AA. The timing of face selectivity and attentional modulation in visual processing. Neuroscience. 2008 Apr; 152(4):1130–1144. https://www.sciencedirect.com/science/article/pii/S0306452208001541, doi: 10.1016/j.neuroscience.2008.01.056.

Papadopouli M, Smyrnakis I, Koniotakis E, Savaglio MA, Brozi C, Psilou E, Palagina G, Smirnakis SM. Brain orchestra under spontaneous conditions: Identifying communication modules from the functional architecture of area V1. bioRxiv. 2024 Mar; https://www.biorxiv.org/content/10.1101/2024.02.29.582364v1, doi: 10.1101/2024.02.29.582364, pages: 2024.02.29.582364 Section: New Results.

Park J, Papoutsi A, Ash RT, Marin MA, Poirazi P, Smirnakis SM. Contribution of apical and basal dendrites to orientation encoding in mouse V1 L2/3 pyramidal neurons. Nature Communications. 2019 Nov; 10(1):5372. https://www.nature.com/articles/s41467-019-13029-0, doi: 10.1038/s41467-019-13029-0, publisher: Nature Publishing Group.

Pasupathy A, Popovkina DV, Kim T. Visual Functions of Primate Area V4. Annual Review of Vision Science. 2020; 6(1):363–385. https://doi.org/10.1146/annurev-vision-030320-041306, doi: 10.1146/annurev-vision-030320-041306, _eprint: https://doi.org/10.1146/annurev-vision-030320-041306.

Petousakis KE, Park J, Papoutsi A, Smirnakis S, Poirazi P. Modeling apical and basal tree contribution to orientation selectivity in a mouse primary visual cortex layer 2/3 pyramidal cell. eLife. 2023 Dec; 12:e91627. https://elifesciences.org/articles/91627, doi: 10.7554/eLife.91627.

Polack PO, Friedman J, Golshani P. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nature Neuroscience. 2013 Sep; 16(9):1331–1339. doi: 10.1038/nn.3464.

Reynolds JH, Chelazzi L. Attentional modulation of visual processing. Annual Review of Neuroscience. 2004; 27:611–647. doi: 10.1146/annurev.neuro.26.041002.131039.

Ringach DL. Spontaneous and driven cortical activity: implications for computation. Current Opinion in Neurobiology. 2009 Aug; 19(4):439–444. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3319344/, doi: 10.1016/j.conb.2009.07.005.

Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG. Signal Timing Across the Macaque Visual System. Journal of Neurophysiology. 1998 Jun; 79(6):3272–3278. https://journals.physiology.org/doi/full/10.1152/jn.1998.79.6.3272, doi: 10.1152/jn.1998.79.6.3272, publisher: American Physiological Society.

Schuman B, Dellal S, Prönneke A, Machold R, Rudy B. Neocortical Layer 1: An Elegant Solution to Top-Down and Bottom-Up Integration. Annual Review of Neuroscience. 2021 Jul; 44(1):221–252. https://www.annualreviews.org/doi/10.1146/annurev-neuro-100520-012117, doi: 10.1146/annurev-neuro-100520-012117.

Semedo JD, Jasper AI, Zandvakili A, Krishna A, Aschner A, Machens CK, Kohn A, Yu BM. Feedforward and feedback interactions between visual cortical areas use different population activity patterns. Nature Communications. 2022 Mar; 13(1):1099. https://www.nature.com/articles/s41467-022-28552-w, doi: 10.1038/s41467-022-28552-w, number: 1 Publisher: Nature Publishing Group.

Semedo JD, Zandvakili A, Machens CK, Yu BM, Kohn A. Cortical Areas Interact through a Communication Subspace. Neuron. 2019 Apr; 102(1):249–259.e4. https://linkinghub.elsevier.com/retrieve/pii/S0896627319300534, doi: 10.1016/j.neuron.2019.01.026.

Serre T, Oliva A, Poggio T. A feedforward architecture accounts for rapid categorization. Proceedings of the National Academy of Sciences. 2007 Apr; 104(15):6424–6429. https://www.pnas.org/doi/10.1073/pnas.0700622104, doi: 10.1073/pnas.0700622104, publisher: Proceedings of the National Academy of Sciences.

Serre T, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust Object Recognition with Cortex-Like Mechanisms. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007 Mar; 29(3):411–426. https://ieeexplore.ieee.org/document/4069258, doi: 10.1109/TPAMI.2007.56, conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

Shen S, Jiang X, Scala F, Fu J, Fahey P, Kobak D, Tan Z, Reimer J, Sinz F, Tolias AS. Distinct organization of two cortico-cortical feedback pathways. bioRxiv; 2020.

Stringer C, Michaelos M, Tsyboulski D, Lindo SE, Pachitariu M. High-precision coding in visual cortex. Cell. 2021 May; 184(10):2767–2778.e15. https://www.sciencedirect.com/science/article/pii/S0092867421003731, doi: 10.1016/j.cell.2021.03.042.

Stringer C, Pachitariu M, Steinmetz N, Carandini M, Harris KD. High-dimensional geometry of population responses in visual cortex. Nature. 2019 Jun; 571(7765):361–365. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6642054/, doi: 10.1038/s41586-019-1346-5.

Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD. Spontaneous behaviors drive multidimensional, brainwide activity. Science. 2019 Apr; 364(6437):eaav7893. https://www.science.org/doi/10.1126/science.aav7893, doi: 10.1126/science.aav7893, publisher: American Association for the Advancement of Science.

Talluri BC, Kang I, Lazere A, Quinn KR, Kaliss N, Yates JL, Butts DA, Nienborg H. Activity in primate visual cortex is minimally driven by spontaneous movements. Nature Neuroscience. 2023 Nov; 26(11):1953–1959. https://www.nature.com/articles/s41593-023-01459-5, doi: 10.1038/s41593-023-01459-5, number: 11 Publisher: Nature Publishing Group.

Tang Y, Gervais C, Moffitt R, Nareddula S, Zimmermann M, Nadew YY, Quinn CJ, Saldarriaga V, Edens P, Chubykin AA. Visual experience induces 4–8 Hz synchrony between V1 and higher-order visual areas. Cell Reports. 2023 Dec; 42(12):113482. https://www.sciencedirect.com/science/article/pii/S2211124723014948, doi: 10.1016/j.celrep.2023.113482.

Wang Q, Burkhalter A. Area map of mouse visual cortex. Journal of Comparative Neurology. 2007; 502(3):339–357. https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.21286, doi: 10.1002/cne.21286, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.21286.

Wosniack ME, Kirchner JH, Chao LY, Zabouri N, Lohmann C, Gjorgjieva J. Adaptation of spontaneous activity in the developing visual cortex. eLife. 2021 Mar; 10:e61619. https://doi.org/10.7554/eLife.61619, doi: 10.7554/eLife.61619, publisher: eLife Sciences Publications, Ltd.

Yamins DLK, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proceedings of the National Academy of Sciences. 2014 Jun; 111(23):8619–8624. http://www.pnas.org/doi/10.1073/pnas.1403112111, doi: 10.1073/pnas.1403112111, publisher: Proceedings of the National Academy of Sciences.

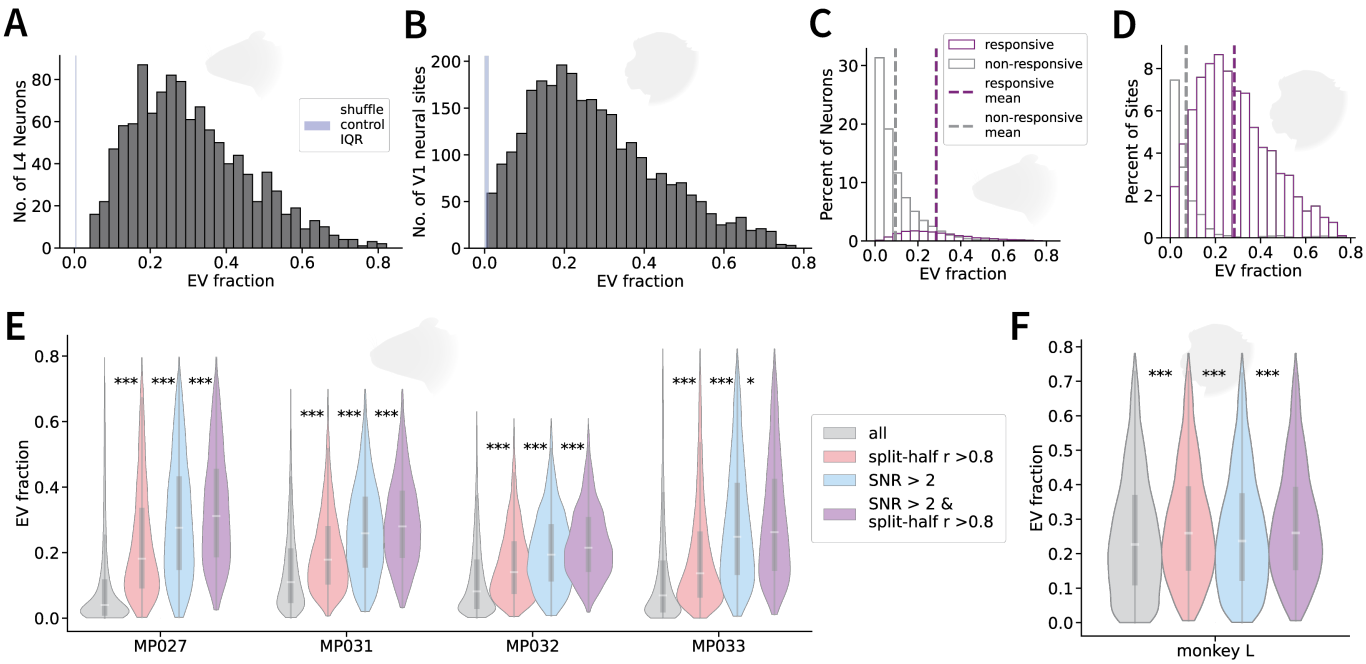

Figure Supplement 1. EV fraction in mouse L4 neurons and macaque V1 neuronal sites and comparison between visual and non-visual neurons/sites. A. Distribution of EV fraction in L2/3→L4 regressions in cells deemed visually responsive in 4 mice and 7 recording days. **B.** Same as A, but for macaque V4→V1 neuronal site regressions. **C** Distribution of visually (purple) and non-visually (gray) responsive neurons in mouse L2/3 and L4. In mouse, we used a conservative criterion to select neurons to be visually responsive, based on an average signal to noise ratio of over 2 and a split-half correlation value of at least 0.8 (more details in Methods). **D**. Same as C but for V1 and V4 sites in macaque. **E.** Differences in EV fractions using filtering methods to determine visually responsive neurons in mouse L2/3 and L4 across the 4 mice. Both a SNR of over 2 along with a split-half correlation value of over 0.8 was used to determine a neuron to be visually responsive. **F**. Same as E, but for the one macaque V1 and V4 sites.

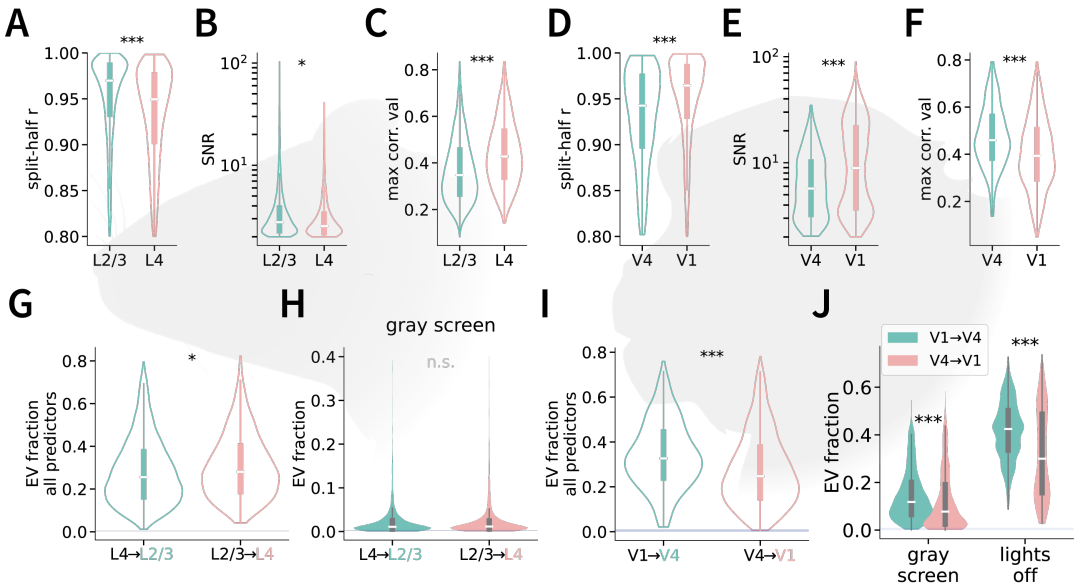

Figure Supplement 2. Neuronal property differences between areas in mouse and monkey. Differences in self-consistency (**A**), SNR (**B**), and max correlation value (**C**) between entire visually responsive neuronal populations in mouse L2/3 and L4 (independent permutation test, here and throughout figure). **D–F** Same as **A–C** but for macaque V4 and V1. **G.** Differences in inter-laminar predictability directions in mouse when using all predictors in their respective L2/3 and L4 layers. **H.** Differences in inter-laminar predictability directions in mouse during gray screen presentation neuronal activity. **I.** Same as **G**, but for macaque V1 and V4. **J.** Differences in inter-cortical predictability between macaque V1 and V4 neuronal activity in response to gray screen presentations and during lights off conditions.

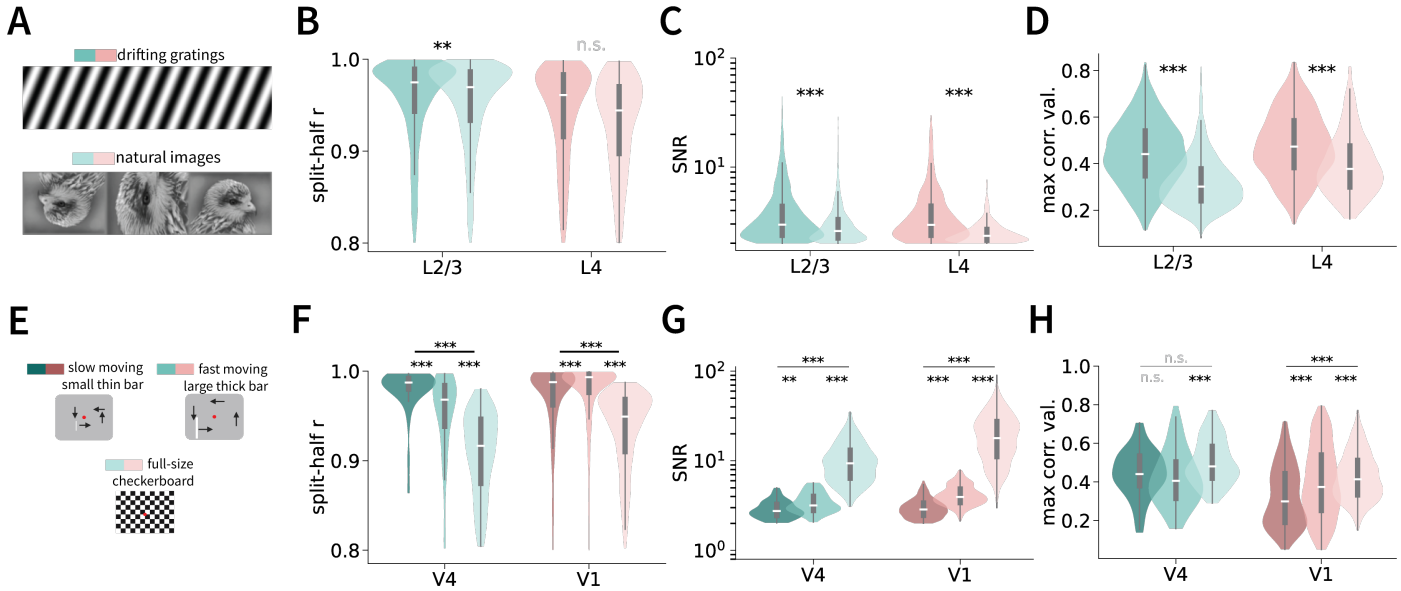

Figure Supplement 3. neuronal activity properties for different stimulus types. A. Sample stimuli for mouse drifting grating and static natural images. **B-D.** Split-half correlation (**B**), SNR (**C**), and maximum correlation values (**D**) for each mouse layer and stimulus type (see color scheme in part **A**. **E.** Sample stimuli for monkeys: full-size checkerboard, slow and fast moving bars. **F-H**. Same as **B-D** for macaque V1 and V4 (see color map for each stimulus condition in **E**.

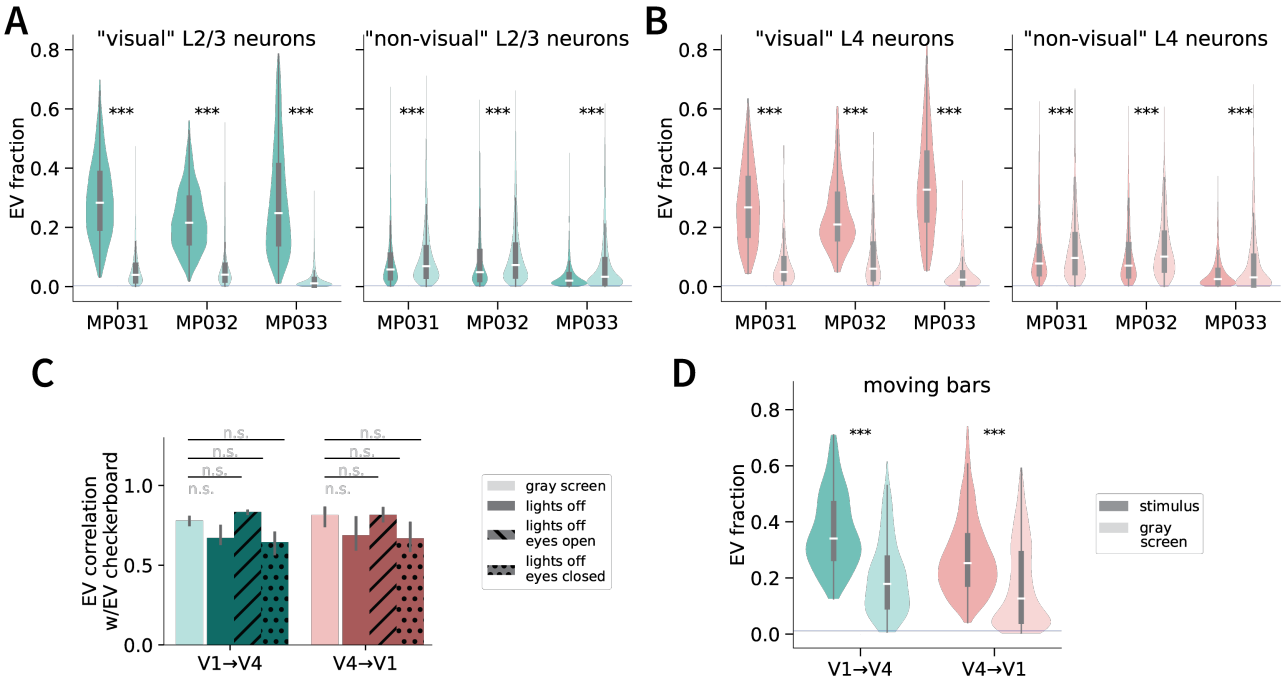

Figure Supplement 4. Comparing stimulus presentation vs. gray screen activity predictions in mouse and macaque. A. Differences in inter-laminar predictability between stimulus presentation and gray screen presentation neuronal activity in L2/3 across the three different mice (MP027 did not undergo gray screen presentation recordings). Left: visually responsive L2/3 neurons. Right: non-visually responsive L2/3 neurons. **B.** Same as **A**, but for mouse L4. **C.** Correlation coefficient values between checkerboard presentation and gray screen and lights off conditions in macaque inter-cortical predictability. **D.** Differences in inter-cortical predictability between moving bar presentation and gray screen activity in macaque V1 and V4 (paired permutation test).

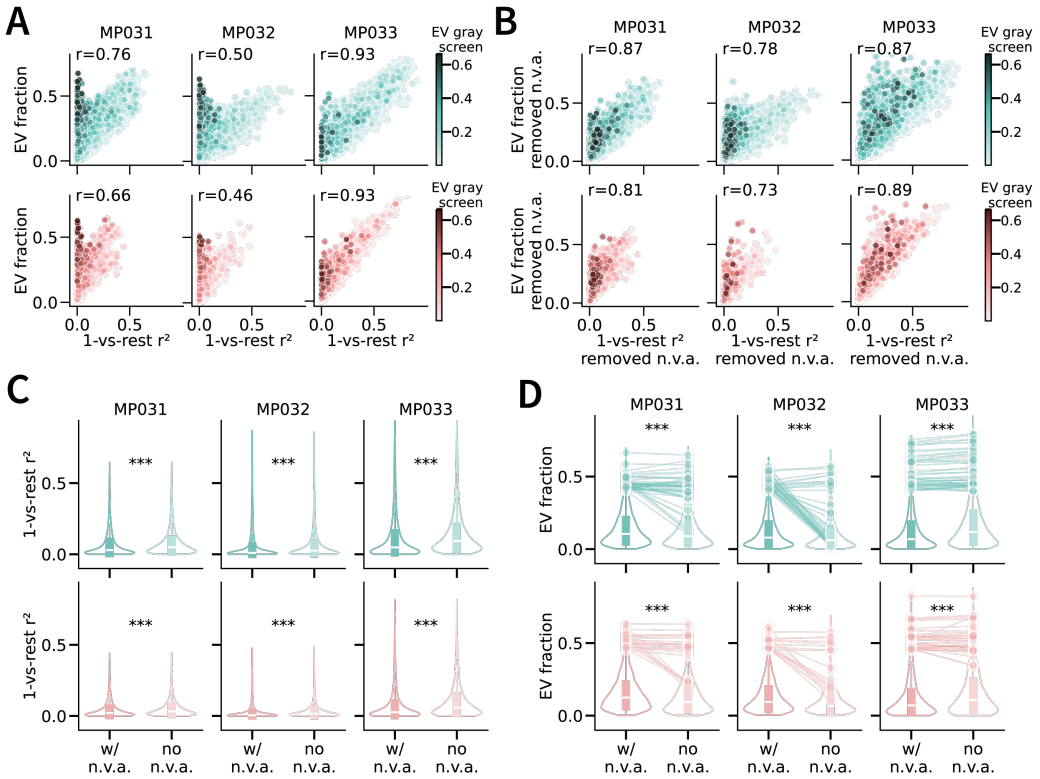

Figure Supplement 5. Bimodal distributions of visual and nonvisually active inter-laminar predictability across mice. A. Relationship between 1-vs-rest squared correlation and EV fraction in L2/3 (top, green) and L4 (bottom, coral) neurons across all mice. **B.** Same relationship between 1-vs-rest squared correlation and EV performanc, but after projecting out "non-visual ongoing" activity (***Stringer et al., 2019a***). **C.** Differences in 1-vs-rest squared correlation values between including and not including non-visual ongoing activity dimensions across three mice (paired permutation test). **D.** Differences in inter-laminar predictability between including and not including non-visual ongoing activity dimensions in L2/3 (top, green) and L4(bottom, coral) across three mice. Sample subset of neurons with initial prediction values of over 0.4 visualized with lineplots.

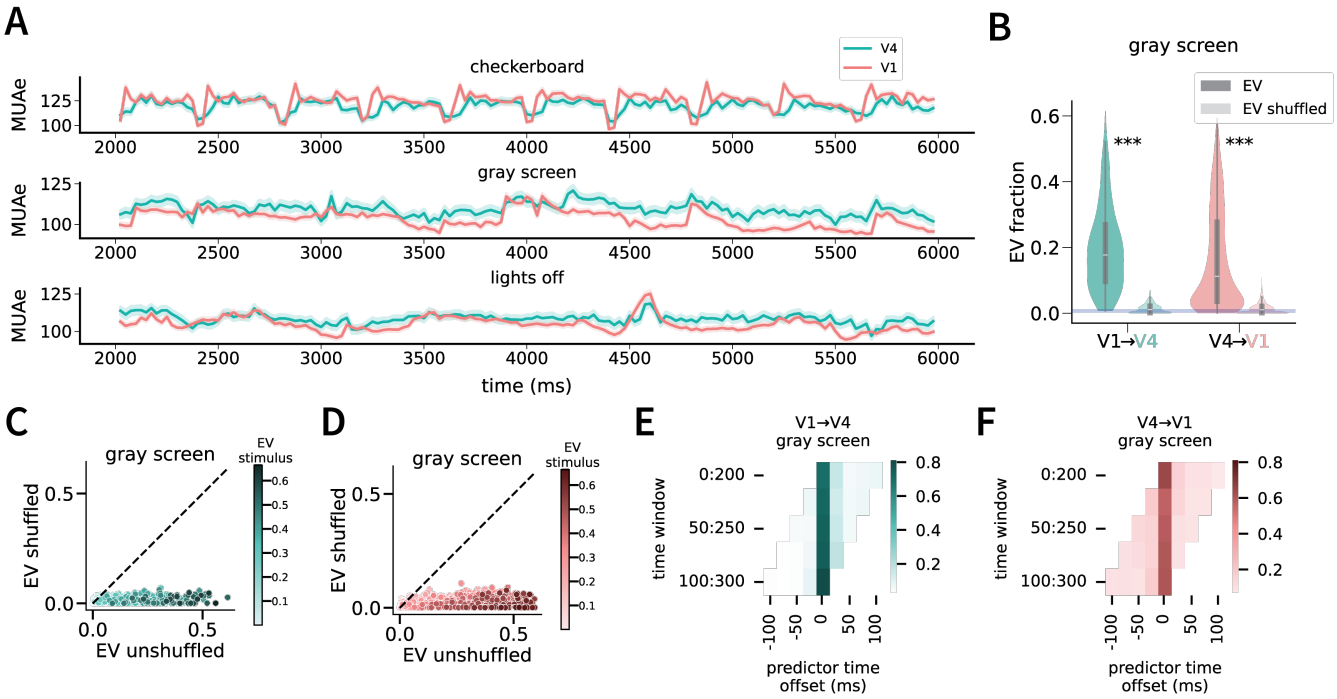

Figure Supplement 6. Predictability across time using gray screen neuronal activity. A. Sample averaged raw MUAe activity in macaque V1 and V4 across the three different conditions. **B.** Differences in EV fraction between unshuffled and shuffled trial-repeat activity during gray screen presentations (paired permutation test). Visualization of relationship between unshuffled and shuffled EV fraction in V4 (**C**) and V1(**D**) during gray screen presentations. **E.** Percentage of neurons with max performance across predictor time offsets in V1→V4 (**E**) and V4→V1 (**F**) directions during gray screen presentations.