

Emergent reliability in sensory cortical coding and inter-area communication

<https://doi.org/10.1038/s41586-022-04724-y>

Received: 29 December 2020

Accepted: 4 April 2022

Published online: 18 May 2022

 Check for updates

Sadegh Ebrahimi^{1,2,3,4}✉, Jérôme Lecoq^{1,2,4,7}, Oleg Rumyantsev^{1,2,5}, Tugce Tasci^{1,2,3}, Yanping Zhang^{1,2,6}, Cristina Irimia^{1,2,4}, Jane Li^{1,4}, Surya Ganguli^{1,5} & Mark J. Schnitzer^{1,2,4,5,6}✉

Reliable sensory discrimination must arise from high-fidelity neural representations and communication between brain areas. However, how neocortical sensory processing overcomes the substantial variability of neuronal sensory responses remains undetermined^{1–6}. Here we imaged neuronal activity in eight neocortical areas concurrently and over five days in mice performing a visual discrimination task, yielding longitudinal recordings of more than 21,000 neurons. Analyses revealed a sequence of events across the neocortex starting from a resting state, to early stages of perception, and through the formation of a task response. At rest, the neocortex had one pattern of functional connections, identified through sets of areas that shared activity cofluctuations^{7,8}. Within about 200 ms after the onset of the sensory stimulus, such connections rearranged, with different areas sharing cofluctuations and task-related information. During this short-lived state (approximately 300 ms duration), both inter-area sensory data transmission and the redundancy of sensory encoding peaked, reflecting a transient increase in correlated fluctuations among task-related neurons. By around 0.5 s after stimulus onset, the visual representation reached a more stable form, the structure of which was robust to the prominent, day-to-day variations in the responses of individual cells. About 1 s into stimulus presentation, a global fluctuation mode conveyed the upcoming response of the mouse to every area examined and was orthogonal to modes carrying sensory data. Overall, the neocortex supports sensory performance through brief elevations in sensory coding redundancy near the start of perception, neural population codes that are robust to cellular variability, and widespread inter-area fluctuation modes that transmit sensory data and task responses in non-interfering channels.

Given a fixed sensory scene or object, sensory recognition is normally reliable. However, sensory cortical neurons have stochastic responses that vary over timescales from seconds to days^{1–4,6,9}. These variations are often shared between cells and across cortical areas^{1–6}, raising basic questions about how neural populations encode and transfer information reliably despite activity fluctuations over multiple spatiotemporal scales^{9–11}.

Many studies have argued that neurons' shared fluctuations constrain the signalling capacity of cortical coding^{3,12–14}, while perhaps also facilitating the decoding of transmitted messages^{6,15,16}. However, the relationships between shared fluctuations, the redundancy of large-scale neural coding and the reliability of sensory cortical representations remain poorly understood. Neural populations can show greater long-term coding stability compared with single cells, but the mechanism for stability and its relationship with shared fluctuations merit further examination^{17–20}.

Human neuroimaging studies usually interpret cofluctuations across brain areas as denoting functional connections for information

transmission^{8,21}. Neuronal recordings have shown that inter-area fluctuations can reflect arousal, neuromodulatory levels or spontaneous movements^{11,22,23} and might also communicate functional information¹⁰. However, whether the cortex uses inter-area fluctuations to encode task-related sensory data has not been tested empirically.

To uncover neural coding and inter-area dynamics promoting reliable sensory processing, we recorded neuronal activity across the entire visual cortex in mice performing a visual task. We analysed thousands of cells, how their visual representations attain coding redundancy and long-term stability, and whether brain areas share information through cofluctuations.

Imaging neuronal activity across the cortex

To study visual processing, we trained head-fixed mice to perform a Go/No-Go task (Fig. 1a, b and Methods). On each trial, mice viewed a moving grating stimulus (duration, 2 s) oriented either horizontally or vertically (termed Go and No-Go stimuli, respectively). Half a second

¹James Clark Center for Biomedical Engineering, Stanford University, Stanford, CA, USA. ²CNC Program, Stanford University, Stanford, CA, USA. ³Department of Electrical Engineering, Stanford University, Stanford, CA, USA. ⁴Department of Biology, Stanford University, Stanford, CA, USA. ⁵Department of Applied Physics, Stanford University, Stanford, CA, USA. ⁶Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. ⁷Present address: Allen Institute, Mindscope Program, Seattle, WA, USA. ✉e-mail: sadegh@stanford.edu; mschnitz@stanford.edu

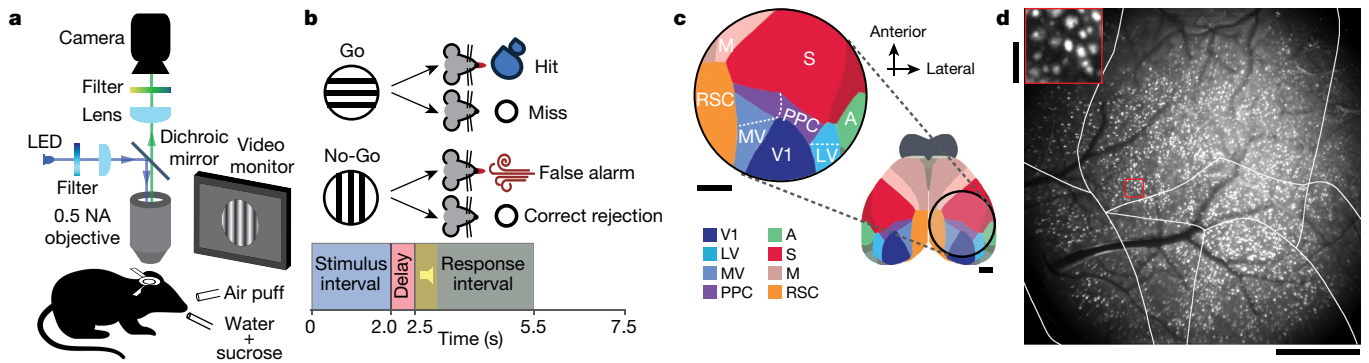


Fig. 1 | Cellular-level imaging across multiple cortical areas during a visual discrimination task. **a**, Imaging of Ca^{2+} activity in thousands of layer 2/3 pyramidal neurons using a custom microscope. **b**, On each trial, mice viewed a moving grating (2 s duration). After a 0.5 s delay, an auditory tone initiated a response period of 3 s, during which mice could respond by licking a spout. Responses to a horizontal grating (the ‘Go’ stimulus) elicited a water reward. If the mouse responded to a vertical grating, it received an air puff and an 8 s timeout before the next trial. Mice performed $83 \pm 3\%$ of the trials correctly

(mean \pm s.e.m.; $n = 6$ mice; Extended Data Fig. 1). **c**, The imaged brain areas. Scale bars, 1 mm. Inset: magnified view. A, auditory cortex; LV, lateral visual cortex; M, motor cortex; MV, medial visual cortex; PPC, posterior parietal cortex; RSC, retrosplenial cortex; S, somatosensory cortex; V1, primary visual cortex. The same abbreviations are used in all subsequent figures. **d**, Maximum projection of a Ca^{2+} video (280 min duration) with 5,292 cells, overlaid with cortical area boundaries. Inset: magnification of the area enclosed within the red box. Scale bars, 1 mm (main image), 0.1 mm (inset).

after the offset of a Go stimulus, the mouse could receive a reward by licking a spout. Incorrect licking after a No-Go stimulus elicited an aversive air puff. To minimize motor-related neural activity during stimulus presentation, we trained the mice to withhold licking until the response period (Fig. 1b). Near the end of training and before brain imaging began, we reduced the grating contrast so mice just surpassed 80% success on both trial types.

As the mice performed the task, we used a fluorescence microscope (16 mm² field of view) to image somatic Ca^{2+} dynamics in neocortical layer 2/3 pyramidal neurons (Fig. 1c, d and Supplementary Video 1). To avoid conflating locomotor-evoked and visual neural signals, we analysed only trials in which locomotion remained $<1 \text{ cm s}^{-1}$. Each recording spanned nearly all of the primary and higher-order visual cortical areas, plus parts of somatosensory, auditory, posterior parietal, motor and retrosplenial cortex. By identifying cells within concatenated datasets, we tracked 21,570 neurons ($3,597 \pm 1,082$ (mean \pm s.d.)) in 6 mice that performed $2,000 \pm 415$ trials over 5–7 days; Figs. 1d and 2a and Extended Data Figs. 1 and 2a–d), thereby attaining long-term and concurrent access to neuronal dynamics in multiple cortical areas.

Variability of cellular-level coding

Across eight cortical areas, many cells preferentially responded to one of the two stimuli, with variable time-dependencies across cells and areas (Extended Data Figs. 2e–h and 3a, b). To characterize cellular coding, we examined correctly performed trials and determined the statistical fidelity, d' , with which one could distinguish the two trial types on the basis of the dynamics of each cell during the stimulus, delay or response intervals. Notably, $(d')^2$ relates to the Fisher information conveyed about trial type^{12–14}. In merged datasets across all days, most cells exhibited tuning to trial type in at least one of the trial periods (16,682 cells with significant tuning; 10,329, 9,204 and 11,958 in the stimulus, delay and response periods, respectively; $P < 0.01$; permutation test; 710–1,340 trials per mouse; Fig. 2b, c and Extended Data Fig. 2h). The fractions of cells tuned to trial type were similar across visual areas, but the distributions of d' varied, especially owing to outlier cells with high d' values (Fig. 2c, d).

Many cells had d' values and coding properties that changed within individual sessions, even while their Ca^{2+} traces retained high signal-to-noise ratios and stable event rates (Extended Data Fig. 1i–k). Some cells increased their d' values while others decreased theirs (Extended Data Fig. 2g, j). These bidirectional changes were balanced

in magnitude, could not result from photobleaching and were unlikely to reflect movement-induced effects because movement generally increases pyramidal cell activity^{11,23,24}.

To assess coding stability, we tested whether cells concentrated their coding responses into subportions of the ~1 h imaging sessions by computing d' separately for the two halves of each session. We also analysed shuffled datasets with random permutations of the trial order. If coding cells concentrate their responses into specific epochs, coding should vary more across half sessions in real than trial-shuffled data, which indeed was the case (Extended Data Fig. 2e), indicative of intrasession coding fluctuations.

Many cells also had variable coding fidelity across days (Extended Data Fig. 2f, h, i). However, as in previous research²⁰, only a minority flipped their coding preference ($1.7 \pm 0.9\%$ of coding cells) and these cells had very low d' values (0.13 ± 0.05 , mean \pm s.d.; $n = 587$ cells that flipped preference in 6 mice). Notably, fluctuations were correlated across timescales; cells with variable intraday coding were about fourfold more likely to have variable across-day coding (Extended Data Fig. 2l). The anatomic comingling of cells with greater and lesser stability (Extended Data Fig. 2i) and correlations between short- and long-term fluctuations make it hard to argue that coding variability arose from imperceptible changes in image quality or focal plane drift.

Time-invariant decoding strategies

Given the non-stationarities in cellular coding, we examined whether an area receiving such variable signals would need to continually adjust its readout strategy to optimally extract stimulus information. Ongoing plasticity might enable such adjustments or, alternatively, neural ensembles might achieve reliability through redundant signalling across multiple cells, information encoded in the correlation structure of neural population activity or combinations thereof^{5,9,14,15,19,25}.

To examine these issues, we trained optimal linear decoders for each brain area to distinguish between the two types of correctly performed trials on the basis of neural ensemble activity in 100 ms time bins (Methods). These ‘instantaneous decoders’ accurately determined the trial type and, as previously³, had a stable form over the latter 1.5 s of the 2 s stimulus presentation (Fig. 3a, b and Extended Data Fig. 3c, f–h). Given this constancy, for the interval 0.5–2 s after stimulus onset, we trained ‘consensus decoders’, the performance of which matched or surpassed that of the instantaneous decoders in most time bins (Extended Data Fig. 3g). Notably, the form of the consensus decoder was

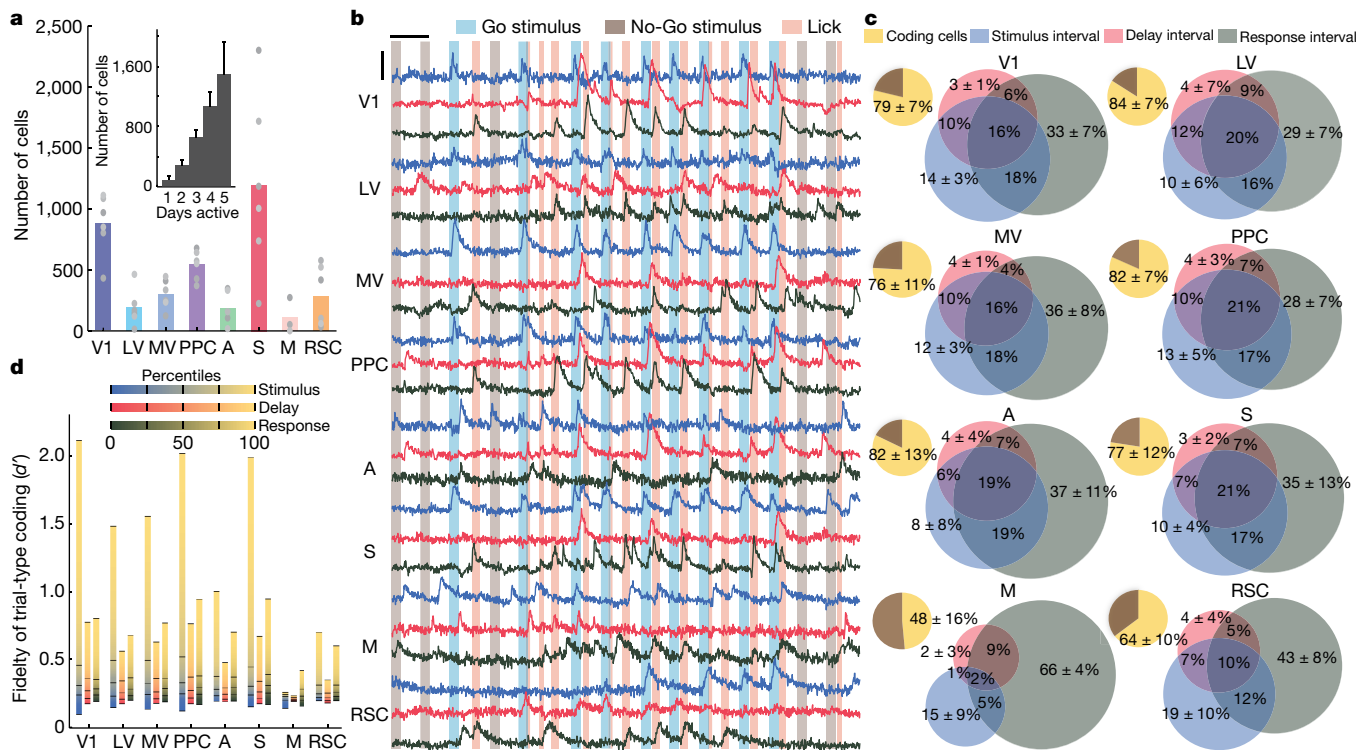


Fig. 2 | Layer 2/3 cells exhibit diverse coding properties during visual discrimination. **a**, The mean numbers of cells identified in each mouse and brain area (total cells, $3,597 \pm 1,082$ (mean \pm s.d.); $n = 6$ mice). Grey points show data from individual mice. Inset: histogram showing the numbers of cells that were active on 1, 2, 3, 4, or all 5 days of the experiment. The error bars (s.d.) were determined as counting errors. **b**, Ca^{2+} traces for three neurons from each of eight areas. Traces of cells responding during stimulus, delay and response intervals are shown in blue, red and black, respectively. Scale bars, 5 s.d.

(vertical), 10 s (horizontal). **c**, Pie charts, percentages of cells in each area significantly encoding the stimulus type (yellow; $P < 0.01$; permutation test; $n = 710\text{--}1,340$ trials) on correct trials, across all sessions. Venn diagrams, the proportions of coding cells with dynamics that significantly encoded the stimulus type during one or more of the intervals within correct trials. Data are mean \pm s.d. over six mice. **d**, For each area, we computed the distribution of cellular d' values for trial-type encoding on correct trials. The plots show d' values for each percentile of the distributions, averaged over six mice.

stable over days (Fig. 3c (inset)), especially for visual areas (Extended Data Fig. 3i (insets)).

This across-day stability led us to train one decoder for each area, plus a separate one for all of the areas grouped together, which we termed ‘common decoders’ and optimized for the 0.5–2 s interval after stimulus onset using all correct trials from all sessions. Notably, common decoders outperformed decoders optimized for single sessions; instead of yielding a suboptimal compromise between the best decoders for different days, common decoders benefited from training on multiple days’ data (Fig. 3c and Extended Data Fig. 3i). However, the existence of successful common decoders stemmed not just from greater training data, for when we trained them on equally sized datasets as single-day decoders the two decoder types performed equivalently (Extended Data Fig. 3l). Although, in principle, common decoders could use stimulus- or choice-related neural activity to discriminate between trial types, in practice, common decoders that were trained on stimulus-period data used only stimulus information (Extended Data Fig. 3j), implying that their stability reflected that of stimulus representations.

To identify a basis for stability, we compared common and single-day decoders using trial-shuffled datasets, in which each cell’s responses were randomly permuted across trials of the same type from the same day (Fig. 3d). Trial shuffling leaves the statistical properties of individual cells unchanged but eradicates correlated fluctuations between cells. In contrast to the results obtained from real data, when we used trial-shuffled data the common decoders performed equivalently or worse than decoders optimized for single days (Fig. 3d). Furthermore, with real datasets, accounting for noise correlations was important for extracting information optimally, as decoders that

ignored noise correlations did much less well, especially for common decoders (Fig. 3e). Overall, accounting for correlated fluctuations in the real data was especially important for constructing decoders that were invariant across days (Extended Data Fig. 3i).

We next examined why accounting for noise correlations was so beneficial to stable decoding performance. Strikingly, in real but not shuffled datasets, day-to-day changes in stimulus-evoked neural responses aligned to the principal eigenvectors of the noise covariance matrix describing trial-to-trial response fluctuations (Fig. 3f and Extended Data Fig. 4a). Mathematical modelling showed that this similarity between fluctuations on distinct timescales allows common decoders to be naturally resistant to both forms of variability, instead of compromising between structures optimized for single days, and that this ‘dual robustness’ emerges even for simple feedforward networks in which activity fluctuations on different timescales propagate through the same pathways (Supplementary Information).

To examine how the mouse’s upcoming responses might have affected stimulus encoding, we trained ‘stimulus-only’ and ‘response-only’ consensus decoders that distinguished either the stimulus or the mouse’s upcoming response, with the other factor held fixed. For example, using trials on which mice withheld licking (No-lick trials), we trained decoders to identify the stimulus type. Cells that made the largest contributions to stimulus- and response-only decoders were interspersed across the cortex (Fig. 3g–j and Extended Data Fig. 4). Stimulus-only decoders attained high accuracy independently of the mouse’s upcoming response ($P < 0.7$; signed-rank test; $n = 6$ mice; Extended Data Figs. 3k and 4), suggesting that the sensory cortex separately encodes stimulus- and choice-related signals. Consistent with this, trial-type decoders for the stimulus period captured stimulus-related information, not

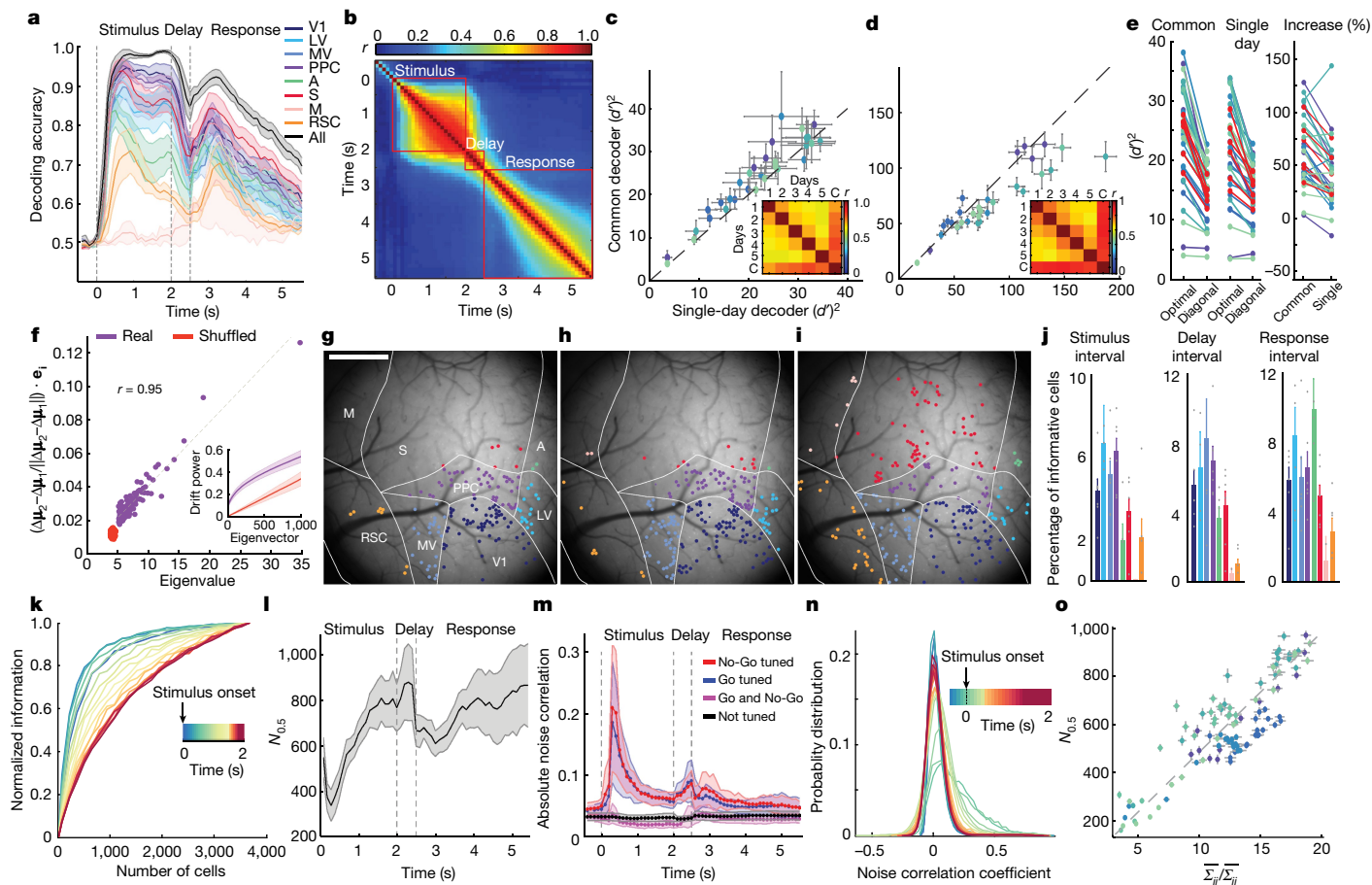


Fig. 3 | See next page for caption.

response-related information. Furthermore, trial-to-trial variations in stimulus encoding were uncorrelated with the mouse’s responses (Extended Data Figs. 3j and 6d), suggesting that incorrect responses were not directly related to the quality of visual coding and instead stemmed from other factors.

Notably, response-only decoders attained significant accuracy during stimulus presentation on Go but not No-Go trials (Extended Data Figs. 3k and 4). Thus, the cortex exhibits signals related to the mouse’s decision or lick preparation on Go trials that are absent on No-Go trials. This may reflect differences in how the brain couples a Go cue to a correct response versus a failure to suppress licking after a No-Go cue. Previous studies have reported similar asymmetries^{26,27}.

Modulation of visual coding redundancy

Since classic studies of motion perception^{5,28}, neuroscientists have appreciated that neural ensembles with correlated fluctuations encode information redundantly, enabling subsets of cells to convey most of the same information as the full ensemble^{3,5,12–14,25}. However, previous research has not directly measured how the redundancy of large-scale neural coding relates to shared fluctuations, especially across brain areas.

We examined three inter-related facets of redundancy: resilience to a hypothetical loss of one cell; the number of cells, $N_{0.5}$, needed to convey 50% of the stimulus-identity information conveyed by all cells; and the levels of correlated fluctuations between cell pairs (Fig. 3k–o and Extended Data Fig. 5). Unexpectedly, correlated fluctuations and visual coding redundancy were time-varying throughout stimulus presentation. Both rose within 100 ms and crested around 200 ms

after stimulus onset, at which time $N_{0.5}$ had its minimum value, stimulus coding was most redundant and correlated fluctuations peaked (Fig. 3k–n). These conditions persisted for only ~300 ms; subsequently, correlated fluctuations and redundancy declined and neurons acted more independently. On average across mice, just after stimulus onset $N_{0.5}$ was about 350 cells, but near stimulation offset $N_{0.5}$ was about 800 cells (Fig. 3l). Within individual mice, the full range of redundancy ($N_{0.5}$) variations was a factor of 3.5 ± 0.5 (mean \pm s.e.m.; $n = 6$ mice).

These changes arose from modulations in task-related neurons. Specifically, correlated fluctuations in similarly tuned stimulus-coding cells rose to a peak around 200 ms after stimulus onset (Fig. 3m). These correlation dynamics had greater amplitudes and distinct kinetics from those of single-cell variability, arose within pairs of cells in the same or different areas and could not be simply explained as due to changes in the activity rates of stimulus-coding cells (Extended Data Fig. 5d–j). Although some cells were modulated by the mouse’s upcoming response ($11 \pm 3\%$ of stimulus-coding cells; mean \pm s.e.m.; $n = 6$ mice; $P < 0.01$, permutation test), response-related modulations had slower kinetics than correlated fluctuations and, at the neural-ensemble level, were orthogonal to stimulus representations and did not affect stimulus-coding redundancy ($N_{0.5}$) (Extended Data Fig. 6c, d). Throughout stimulus presentation, $N_{0.5}$ varied inversely with correlated noise levels in similarly tuned cell pairs, with the same proportionality in all mice ($r = 0.9$; $P < 1.4 \times 10^{-25}$; Fig. 3o). Thus, the 3.5-fold variations in coding redundancy seen in individual mice reflected approximately comparable variations in correlated noise among task-related neurons. Since correlated fluctuations probably arise from the shared inputs of cells^{3,29}, the invariant proportionality constant probably reflects invariant aspects of mouse cortical connectivity. Overall, in contrast to studies that assessed widespread noise

Fig. 3 | Accounting for correlated fluctuations among task-related cells facilitates stable representations of stimulus type. **a**, Mean accuracies for inferring stimulus identity using optimal instantaneous (100 ms time bins) linear decoders of activity for individual (coloured traces) or all brain areas (black trace). Dashed lines in **a**, **l**, **m** demarcate stimulus, delay and response intervals. Data are mean \pm s.e.m. across six mice **b**, Mean similarities between all pairs of instantaneous decoders, assessed by correlation coefficients between pairs of decoder weights for all cells in each mouse ($n = 6$ mice). Given the decoder constancy across stimulus presentation, in **c–j** we trained ‘consensus’ decoders, optimized for 0.5–2.0 s after the stimulus onset (Extended Data Fig. 3f, h). **c**, **d**, To assess decoder stability, we trained ‘common’ consensus decoders on data from all days and compared them to consensus decoders trained on data from single days. We evaluated real (**c**) and trial-shuffled (**d**) datasets in which each cell’s Ca^{2+} traces were randomly permuted across trials of the same stimulus type from the same day. Each blue shade in **c–e** denotes data from one mouse during stimulus presentation. Each data point in **c** and **d** is from one session and shows the stimulus-identity information $(d')^2$ conveyed by common and single-day decoders given identical test datasets from individual days. On real datasets, common decoders outperformed single-day decoders (**c**). On trial-shuffled datasets, single-day decoders generally outperformed common decoders (**d**). Data are mean \pm s.d. across 100 random divisions of each dataset into thirds, for dimensionality reduction, decoder training and testing. Insets: correlation coefficient, r , values between consensus decoders from individual days and the common decoder (C), averaged over six mice (Extended Data Fig. 3i). **e**, Left, optimal linear decoders outperformed diagonal decoders that ignore correlated fluctuations ($68 \pm 6\%$ (mean \pm s.e.m.; $P < 1.7 \times 10^{-6}$) and $40 \pm 5\%$ ($P < 2.3 \times 10^{-6}$) more information was captured by the optimal decoders of trial type for common and single-day decoders of activity during stimulus presentation; signed-rank test; $n = 30$ sessions in 6 mice). Right, the superiority of optimal over diagonal decoders was greater for common than single-day decoders. Increases in $(d')^2$ for optimal versus diagonal decoders were $55 \pm 26\%$ (mean \pm s.e.m.) greater for common than single-day decoders ($P < 4.9 \times 10^{-5}$, signed-rank test; $n = 30$ sessions). Each connected pair of blue-shaded points shows results from one session and one mouse; red points show mean values for individual mice. **f**, Day-to-day drifts in neural responses were aligned with within-day, trial-to-trial fluctuations. To assess day-to-day drift, we computed the unity-normalized vector between the mean neural ensemble responses to each stimulus on consecutive days, $(\mu_2 - \mu_1)/(\|\mu_2 - \mu_1\|)$. To characterize trial-to-trial

fluctuations, we computed the noise covariance matrix of ensemble responses, averaged over both stimuli, for the first day of all consecutive pairs of days. We projected $(\mu_2 - \mu_1)/(\|\mu_2 - \mu_1\|)$ onto this matrix’s eigenvectors, $\{e_i\}$, and averaged over both stimuli and all pairs of consecutive days. Day-to-day drifts aligned with within-day, principal noise eigenvectors in real (purple points; $r = 0.95$; $P < 10^{-50}$) but not trial-shuffled (red points; $r = 0.02$; $P = 0.82$) data. Inset: cumulative plots of the fraction of the power of day-to-day variations within the subspace defined by the first n noise eigenvectors (where n is the abscissa value) for real (purple) and trial-shuffled (red) data. **g–j**, Cells contributing most to the performance of stimulus-only decoders were interspersed across the cortex. Maps of these most-informative cells (with decoder weights that deviated >2 s.d. from the mean) are shown for one mouse (stimulus interval (**g**), delay interval (**h**) and response interval (**i**)) averaged over both response types. Scale bar, 1 mm. **j**, Mean \pm s.e.m. (six mice) percentages of the most-informative cells in each brain area (colours as in **a**). For response decoder results, see Extended Data Fig. 4h–m. **k**, Coding redundancy peaked just after stimulus onset. For each time bin after stimulus onset (colours), we used instantaneous decoders to measure the information $(d')^2$ conveyed about stimulus identity by subsets of cells randomly chosen across all areas. Data are from one mouse and are averages over 100 subsets of each size, normalized to the result for all cells. For results for all mice and delay and response periods, see Extended Data Fig. 5b, c. s.e.m. values (not shown) are $<8\%$ for all points. **l**, Mean ensemble sizes, $N_{0.5}$, at which $(d')^2$ reached its half-maximum, estimated for each time bin using instantaneous decoders of activity across all imaged areas. Data are mean \pm s.e.m. across six mice. **m**, The absolute values of mean noise correlations in Ca^{2+} event rates for pairs of the most-informative cells both tuned to Go stimuli (blue trace), both tuned to No-Go stimuli (red trace) or oppositely tuned (magenta trace). Black trace, results for untuned cells. Data are mean \pm s.e.m. across six mice. **n**, Cell pairs with similar stimulus tuning had their greatest noise correlation coefficients just after stimulus onset. Data show distributions of these coefficients at different times (denoted in colours) pooled over six mice. Error bars (s.d.) are too small to be visible. **o**, $N_{0.5}$ versus the ratio of the mean of the noise covariance matrix’s diagonal elements (Σ_{ii}) to the mean of its non-diagonal elements (Σ_{ij}) for the most-informative neurons. Each data point is from one mouse and time bin during stimulus presentation. The colours denote individual mice and reveal a linear relationship ($r = 0.9$; $P < 1.4 \times 10^{-25}$), consistent with mice having statistically similar neural connectivity matrices. Data are mean \pm s.e.m. over 100 subsamplings of cells (y-axis) or 51–296 cells (x-axis).

correlations with a lower time-resolution¹¹, during passive viewing^{3,10,11}, or without cellular resolution²³, here noise correlations in task-related neurons increased during early phases of perception to more than triple the redundancy of sensory encoding.

We next examined how much of the information, $(d')^2$, provided by our decoders was redundant across brain areas. Decoder outputs proved to be highly correlated between sensory areas; if on one trial stimulus encoding in one area was weaker or stronger than average, this was usually so in other areas (Fig. 4a–c and Extended Data Fig. 6). This interdependence and the resulting coding redundancy across areas had a similar time-dependence to the noise correlations among task-related cells. Within ~ 200 ms of stimulus onset, decoder score correlations peaked, yielding approximately a threefold redundancy across the brain areas examined (Fig. 4d). This was not just from replication of information within the primary visual cortex (V1), because the full set of cells conveyed almost twice the information of those in V1 (Extended Data Fig. 4b), suggesting that higher-order areas receive additional information from outside V1. After attaining their peak values, coding redundancy and decoder score correlations declined for the remainder of visual stimulation. Near stimulus offset, visual representations in different areas were almost mutually independent, consistent with the vanishing correlated noise levels between cell pairs (Figs. 3m and 4d). Overall, time-varying cofluctuations among task-related cells greatly impacted visual processing, leading to severalfold increases in coding resilience (Extended Data Fig. 5i), redundancy and inter-area correlations that peaked soon after stimulus onset.

Communication through inter-area fluctuations

Activity cofluctuations of cell ensembles are thought to reflect shared connectivity, such as common inputs, or direct interconnections^{10,30,31}. In the absence of sensory stimuli, such fluctuations can reflect an animal’s spontaneous behaviour¹¹. During sensory tasks, previous studies examined shared fluctuations across pairs of electrodes^{32–35} and decoder score correlations across a pair of brain areas³⁶, but the anatomic distributions and time-dependencies of neuronal cofluctuations across multiple areas and how they relate to task performance remain unexplored¹⁰.

To identify cofluctuating cell ensembles across pairs of areas, we applied canonical correlation analysis (CCA) to mean-subtracted neural activity traces, which represent trial-by-trial activity fluctuations. CCA identifies dimensions of shared activity and paired sets of dynamical or communication modes¹⁰ (CCA modes) ranked by their levels of co-varying activity (Extended Data Figs. 7–9 and Methods). During visual stimulation, the number of CCA modes with significant cofluctuations varied across different pairs of areas but was generally <20 in our datasets (Extended Data Fig. 7). Inter-area, CCA fluctuation modes comprised about 60% of the total power of all cortical fluctuations, implying a majority of fluctuation power during visual stimulation propagates across cortical regions (Fig. 4e, f).

Given the time dependence of the correlated fluctuations of task-related cells, we compared the CCA modes arising during visual stimulation to those present just beforehand. Notably, by around 200 ms after stimulus onset, CCA modes present in inter-trial intervals

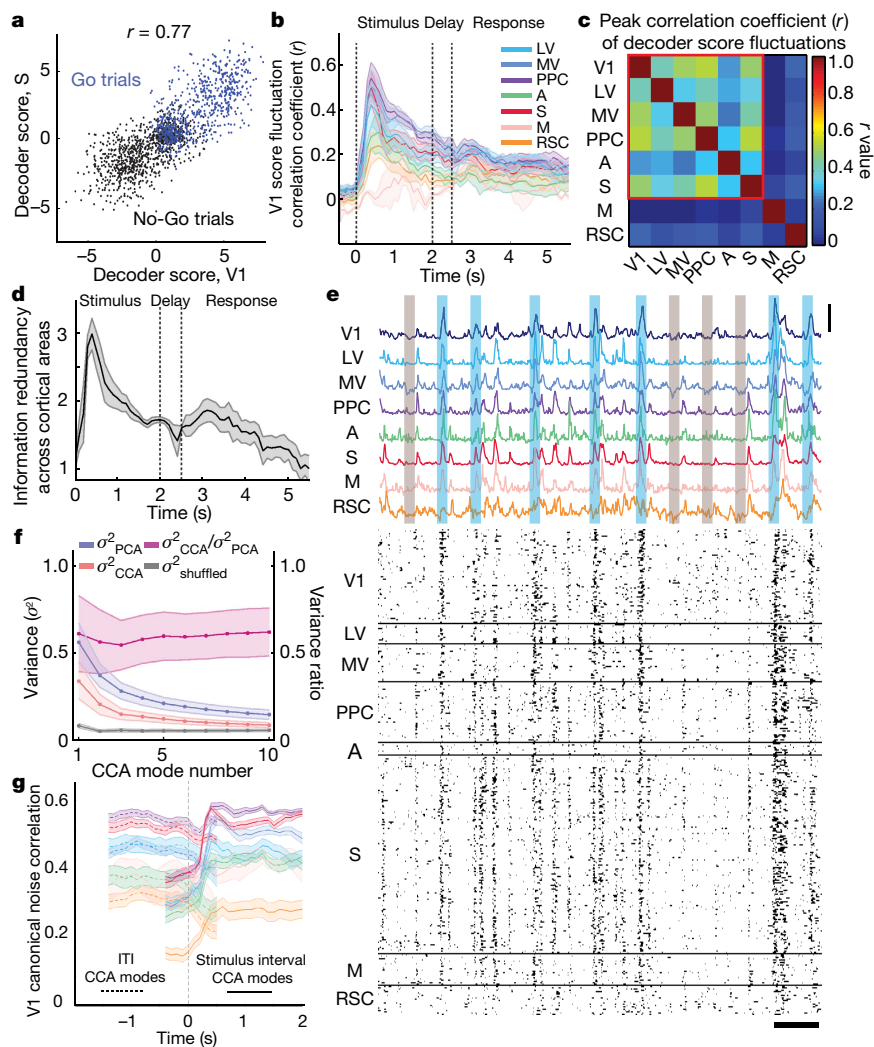


Fig. 4 | Inter-area fluctuations and stimulus encoding redundancy peaked around 200 ms after stimulus onset. **a**, Different sensory areas had strongly correlated decoder scores. Shown is an example scatter plot for one mouse for which we trained stimulus-type decoders using either V1 or S1 activity from 0.5–0.6 s after stimulus onset on correctly performed trials. Each data point shows the two decoder scores on one trial. Extended Data Fig. 6a shows another example from areas PPC and RSC. **b**, **c**, Correlation coefficients, r , for decoder scores peaked \sim 200 ms after stimulus onset. **b**, Time-varying mean \pm s.e.m. ($n = 6$ mice) r values between V1 and seven other regions. **c**, Peak r values for all area pairs, averaged over mice (Extended Data Fig. 6b, d). Dashed lines in **b** and **d** demarcate the stimulus, delay and response periods. **d**, Redundancy of stimulus encoding across the cortex peaked \sim 200 ms after stimulus onset and then declined towards unity. Data are mean \pm s.e.m. over six mice. **e**, Bottom, raster plots of Ca^{2+} events in individual cells (from eight areas in one mouse) with large contributions to inter-area cofluctuation modes found by CCA. Top, coloured traces show the dynamics of the largest CCA modes between V1 and seven other areas. The V1 trace is an average of results from all seven analyses. Cyan and grey shading mark Go and No-Go stimulus presentations, respectively. Scale bars, 5 s.d. (vertical), 10 s (horizontal). **f**, Inter-area cofluctuations comprised about 60% of the total power of cortical noise modes. The plot shows the mean powers, σ^2 (the variance), of the ten largest CCA modes (red curve, left y-axis), averaged over all 28 area pairs and both areas per pair, and the mean power of the 10 largest noise modes (blue curve) found by principal component analysis (PCA) of fluctuations in each area, averaged over all 8 areas. Noise modes found by randomly shuffling weights from CCA (black curve) had substantially less power. Ratios of noise power in CCA and PCA modes (magenta curve, right y-axis) were consistently around 60%. For **f** and **g**, data are mean \pm s.e.m. over 6 mice. **g**, Distinct inter-area cofluctuations arose during visual stimulation and ITIs, defined as 2 s intervals preceding stimulus onsets. We separately applied CCA to ITIs and stimulus presentation periods. The time-varying correlation coefficients for the largest noise modes between V1 and seven other areas are plotted (coloured as in **b**). At the stimulus onset, correlated activity rose considerably in modes found during visual stimulation, whereas activity in the ITI modes declined (Extended Data Fig. 8).

(ITIs) had decayed and a new set of modes had activated (Fig. 4g and Extended Data Fig. 8). Thus, inter-area fluctuations in animals nominally at rest^{11,37} appear to be distinct from those during an active sensory task.

To characterize the spatial structure of inter-area fluctuations, for each choice of brain area as a source, we quantified the similarity of its CCA modes with each of the seven other imaged areas. Strikingly, for every source area, the primary communication mode was nearly the same, irrespective of the target, implying that there was a global mode of cofluctuations (Fig. 5a, b). Secondary modes were more localized and shared across subsets of areas. For example, V1 shared one

secondary mode with the auditory cortex and the somatosensory cortex, and another with the lateral visual and medial visual cortical areas and the posterior parietal cortex (Fig. 5a–c). Thus, CCA revealed a hierarchical structure in which each area shared a global fluctuation mode with all other areas, and distinct secondary modes with different sets of areas.

We examined whether cofluctuation modes carried signals relating to the visual discrimination task (Fig. 5d, e). About 0.5 s after stimulus onset, activity in the second and higher CCA modes accurately encoded stimulus identity. Up to \sim 80% of the total information encoded in the

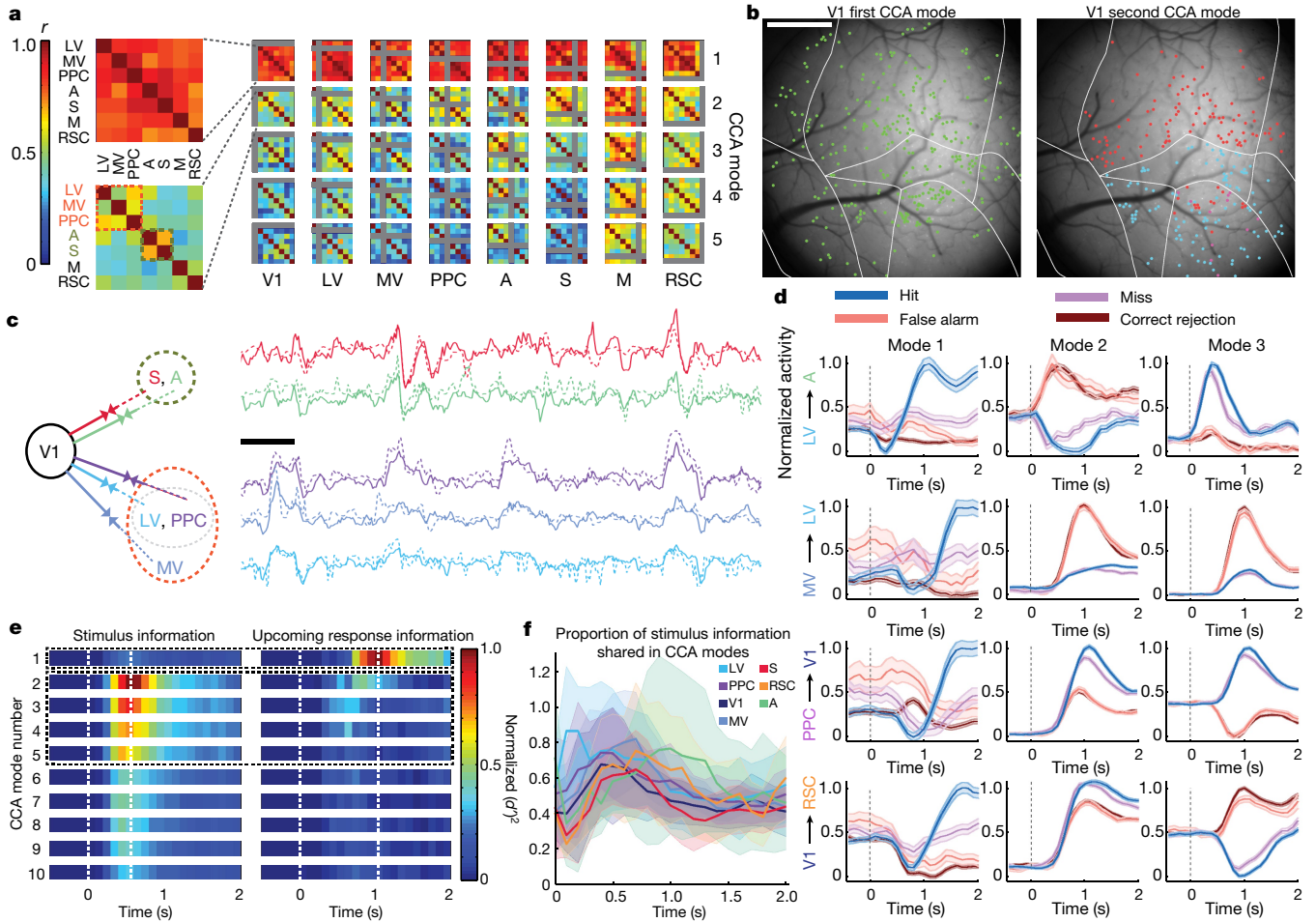


Fig. 5 | Orthogonal inter-area cofluctuations communicate sensory data and the mouse's upcoming response.

a, Each matrix shows correlation coefficients, r , for CCA modes between one of eight source areas (listed at bottom) and two target regions (arranged as in the insets). A large matrix element value indicates that the source cofluctuated with the two targets using a similar activity mode; small values imply distinct cofluctuation modes. Results are shown for the five largest CCA modes for each source–target pair, averaged over six mice. The largest CCA mode (top row) was largely invariant to source–target choices and thus globally shared across areas (the mean r values of the largest modes for individual mice were 0.99, 0.95, 0.85, 0.91, 0.92 and 0.68). Insets: magnified views for the largest CCA modes involving V1 and one of seven other areas (top), and the second-largest modes between V1 and these other areas (bottom). In five out of six mice, there were at least two clusters (orange and olive fonts) of secondary modes with moderate similarity (schematic in **c**). Modes involving V1 and LV, MV or PPC comprised one cluster; modes involving V1 and either area A or S comprised another. **b**, Left, map of neurons (green) contributing significantly (weights deviating by >2 s.d. from mean values) to the global fluctuation mode in one mouse. Right, map of neurons in the two clusters of second-largest CCA modes involving V1 (see **a** and **c**). Cells marked red contributed to cofluctuations between V1 and either S or A. Cells marked cyan contributed to cofluctuations between V1 and LV, MV or PPC. **c**, Left, clustering revealed two subsets of target areas with similar second-largest CCA modes in V1, as seen in **a** and **b**. Right, ten example activity traces for these modes, coloured to match the areas on the left. Solid traces show activity within the CCA mode in V1. Dotted traces show activity in the target area's CCA mode. Scale bar, 5 s. **d**, Aggregate neural Ca^{2+} signals in one

mouse within the population vector dimensions determined by the largest three CCA modes (columns), for four different area pairs (rows) and trial outcomes (coloured traces). The dashed line marks the stimulus onset. Ordinate values are shifted and normalized to lie within [0,1]. Data are mean \pm s.e.m. $n = 100$ –678 trials. **e**, Right, the global fluctuation mode, identified in **a**, lies in the dimension encoding information late in the stimulus period about the mouse's upcoming response. Left, the second- to fifth-largest CCA modes lie in dimensions encoding stimulus type. The results are from a CCA analysis of V1, LV, MV, PPC, A and S in which the cell ensembles significantly encoded stimulus type or the mouse's upcoming response ($P < 0.01$, permutation test across trials of different types, using equal trials of each type; $n = 52$ –854 trials per type per mouse). We analysed the 15 area pairs, projected activity in each area onto the dimensions identified, and computed how much information, $(d')^2$, this activity subset encoded about the stimulus type (on Lick and No-Lick trials, on which the mouse, respectively, did or did not exhibit a licking response during the response interval), or about the upcoming response (on Go trials). The time-varying $(d')^2$ values, averaged over both projections for each of 15 area pairs in 6 mice, are shown for the 10 largest CCA modes (Extended Data Fig. 8). **f**, To determine the proportion of stimulus information shared through CCA modes, we plotted the total information encoded in CCA modes between a source (coloured traces) and the other seven areas, relative to the total information encoded within the source. Visual areas had a preponderance of their stimulus information encoded within CCA modes, especially early during stimulus presentation. Ratios for non-visual areas peaked later in the trial. Data are mean \pm s.e.m. over six mice (Extended Data Fig. 9a).

cortex about stimuli identity was shared between areas in these modes, which conveyed almost nothing about the mouse's upcoming response (Fig. 5e–f and Extended Data Fig. 9a). Later, around 1 s into stimulus presentation, on Go trials the global cofluctuation mode encoded the upcoming response but no stimulus information, consistent with

our ability to decode upcoming responses on Go but not No-Go trials. Overall, the neocortex uses non-interfering communication channels, that is, orthogonal cofluctuation modes, to convey stimulus- and response-related signals to distinct sets of areas, in a targeted and global manner, respectively.

Discussion

By tracking neurons across all visual cortical areas, our study reveals information-processing mechanisms that are likely to underlie reliable sensory performance. Historically, neuroscientists viewed correlated neuronal fluctuations as imposing limits on coding accuracy^{5,12–14}, which our study supports. However, our data also show that accounting for correlated fluctuations facilitates the long-term reliability of neural population activity decoders, because day-to-day variations in population coding strongly correlate with the faster coding variations occurring within individual days. This similarity across timescales arises even in simple network models and enables decoding strategies that are intrinsically robust to both forms of variability (Supplementary Information). Decoders that neglect correlated fluctuations lack this dual robustness.

Beginning <100 ms and reaching an apex around 200 ms after stimulus onset, task-related neurons across the cortex momentarily increase their correlated fluctuations for ~300 ms. Importantly, these rapid dynamics in no way conflict with reports that variability in the activity of individual cells declines after stimulus onset³⁸, a pattern that our data confirm (Extended Data Fig. 5e–g). Moreover, the modulation of shared fluctuations seen here in mice performing a visual task contrasts with findings in untrained mice passively viewing stimuli, during which modulations of shared fluctuations were unapparent in V1³. Thus, task performance, long-term training or both might alter the dynamics of correlated fluctuations^{19,39}.

The stimulus-evoked increase in shared fluctuations among task-related cells boosts the redundancy of cortical representations severalfold within a ~300 ms interval. The transient, shared fluctuation modes convey a majority (about 80%) of sensory information across cortical areas within signalling streams orthogonal to that conveying the animal's response. Here, information about the mouse's upcoming response arose in a unique, global mode of fluctuations starting around 0.6 s and peaking about 1 s after stimulus onset. In visual tasks without a delay period, choice-related fluctuations arose sooner after stimulus onset^{40,41}.

In our experiments, the time-interval following the redundancy peak, namely 0.5–2 s after stimulus onset, was when our stimulus decoders attained a stable form (Fig. 3b). Our analyses of long-term decoder stability used data from this 0.5–2 s interval and showed that common decoders can succeed across days without need for daily adjustments. However, these results carry no implications regarding the long-term stability of stimulus decoders trained on time bins within the 0–0.5 s interval, during which decoder forms were changing too rapidly for us to draw conclusions about long-term stability.

The rise and decay of shared fluctuations seen here after stimulus onset may reflect successive feedforward and feedback phases of information flow across sensory cortical areas^{42–44}. In this view, the early sensory cortex uses redundant, inbound sensory data to represent the basic features of a stimulus within the first few hundred milliseconds of its appearance; during later sensory processing, probably involving feedback from higher-order areas, the representations become less redundant and more efficient. This transition, which probably occurs more quickly in primates than mice, may reflect a shift in spiking patterns from those driven initially mainly by incoming sensory signals, arriving through overlapping connections, to those reflecting a rising influence of top-down or recurrent signals propagating through distinct circuitry. This processing shift may help to relate local visual features to their global context or task demands^{42–44}.

The time-varying, anatomic patterns of shared fluctuations are likely to support inter-area communication within distinct subnetworks. Human neuroimaging studies describe a 'default-mode' network of areas with characteristic co-fluctuations that typify the brain's resting state⁷, and other sets of functionally connected areas that co-fluctuate during performance of specific tasks²¹. Here, inter-area

co-fluctuations during a visual task differed from those during ITIs, providing cellular-level evidence of task-dependent changes in the brain's functional connectivity. Bolstering the idea that shared fluctuations subserved specific components of animal behaviour, information about sensory stimuli and upcoming responses was communicated to distinct groups of areas, in orthogonal fluctuation modes and with distinct timing. Future work should quantify the extent to which fluctuation modes are task specific or generalize across tasks with similar components.

It is striking that response-related data were transmitted within a global fluctuation mode that engaged every area examined. Past observations of widespread fluctuations came from animals with no active task to perform^{10,11} or in which fluctuations reflected spontaneous movements or arousal²³. Notably, widespread dissemination of perceptual decisions across brain areas distinguishes some models of conscious perception⁴⁵ and, when related to reward expectation, is a key element in some models of reinforcement learning⁴⁶. As previous reports suggest that brain connectivity might resemble 'small-world' networks^{47,48}, we simulated small-world networks with varying connectivity and linear dynamical fluctuations, but they all lacked a global fluctuation mode; however, networks in which a single source broadcasted common signals to multiple areas did exhibit a global mode (Extended Data Fig. 9). Future research should determine whether such a broadcast exists in the mammalian brain and, if so, in which area or areas it originates.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04724-y>.

- Faisal, A. A., Selen, L. P. & Wolpert, D. M. Noise in the nervous system. *Nat. Rev. Neurosci.* **9**, 292–303 (2008).
- Lutcke, H., Margolis, D. J. & Helmchen, F. Steady or changing? Long-term monitoring of neuronal population activity. *Trends Neurosci.* **36**, 375–384 (2013).
- Rumyantsev, O. I. et al. Fundamental bounds on the fidelity of sensory cortical coding. *Nature* **580**, 100–105 (2020).
- Stein, R. B., Gossen, E. R. & Jones, K. E. Neuronal variability: noise or part of the signal? *Nat. Rev. Neurosci.* **6**, 389–397 (2005).
- Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).
- Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N. & Harvey, C. D. Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell* **170**, 986–999 (2017).
- Greicius, M. D., Supekar, K., Menon, V. & Dougherty, R. F. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* **19**, 72–78 (2009).
- Rosenberg, M. D. et al. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat. Neurosci.* **19**, 165–171 (2016).
- Montijn, J. S., Meijer, G. T., Lansink, C. S. & Pennartz, C. M. Population-level neural codes are robust to single-neuron variability from a multidimensional coding perspective. *Cell Rep.* **16**, 2486–2498 (2016).
- Semedo, J. D., Zandvakili, A., Machens, C. K., Byron, M. Y. & Kohn, A. Cortical areas interact through a communication subspace. *Neuron* **102**, 249–259 (2019).
- Stringer, C. et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, 255 (2019).
- Abbott, L. F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
- Averbeck, B. B. & Lee, D. Effects of noise correlations on information encoding and decoding. *J. Neurophysiol.* **95**, 3633–3644 (2006).
- Moreno-Bote, R. et al. Information-limiting correlations. *Nat. Neurosci.* **17**, 1410–1417 (2014).
- Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A. & Yuste, R. Controlling visually guided behavior by holographic recalling of cortical ensembles. *Cell* **178**, 447–457 (2019).
- Graf, A. B., Kohn, A., Jazayeri, M. & Movshon, J. A. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat. Neurosci.* **14**, 239–245 (2011).
- Ziv, Y. et al. Long-term dynamics of CA1 hippocampal place codes. *Nat. Neurosci.* **16**, 264–266 (2013).
- Xia, J., Marks, T. D., Goard, M. J. & Wessel, R. Stable representation of a naturalistic movie emerges from episodic activity with gain variability. *Nat. Commun.* **12**, 5170 (2021).
- Gonzalez, W. G., Zhang, H., Harutyunyan, A. & Lois, C. Persistence of neuronal representations through time and damage in the hippocampus. *Science* **365**, 821–825 (2019).

20. Deitch, D., Rubin, A. & Ziv, Y. Representational drift in the mouse visual cortex. *Curr. Biol.* **31**, 4327–4339 (2021).
21. Sridharan, D., Levitin, D. J. & Menon, V. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl Acad. Sci. USA* **105**, 12569–12574 (2008).
22. Allen, W. E. et al. Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science* **364**, 253 (2019).
23. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* **22**, 1677–1686 (2019).
24. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).
25. Montani, F., Kohn, A., Smith, M. A. & Schultz, S. R. The role of correlations in direction and contrast coding in the primary visual cortex. *J. Neurosci.* **27**, 2338–2348 (2007).
26. Goard, M. J., Pho, G. N., Woodson, J. & Sur, M. Distinct roles of visual, parietal, and frontal motor cortices in memory-guided sensorimotor decisions. *eLife* **5**, e13764 (2016).
27. Poort, J. et al. Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* **86**, 1478–1490 (2015).
28. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
29. Kanitscheider, I., Coen-Cagli, R. & Pouget, A. Origin of information-limiting noise correlations. *Proc. Natl Acad. Sci. USA* **112**, E6973–E6982 (2015).
30. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
31. Yu, Y., Stirman, J. N., Dorsett, C. R. & Smith, S. L. Mesoscale correlation structure with single cell resolution during visual coding. Preprint at *bioRxiv* <https://doi.org/10.1101/469114> (2018).
32. Gregoriou, G. G., Gotts, S. J. & Desimone, R. Cell-type-specific synchronization of neural activity in FEF with V4 during attention. *Neuron* **73**, 581–594 (2012).
33. Gregoriou, G. G., Gotts, S. J., Zhou, H. & Desimone, R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207–1210 (2009).
34. Ruff, D. A. & Cohen, M. R. Attention increases spike count correlations between visual cortical areas. *J. Neurosci.* **36**, 7523–7534 (2016).
35. van Kempen, J. et al. Top-down coordination of local cortical state during selective attention. *Neuron* **109**, 894–904 (2021).
36. Chen, J. L., Voigt, F. F., Javadzadeh, M., Krueppel, R. & Helmchen, F. Long-range population dynamics of anatomically defined neocortical networks. *eLife* **5**, e14679 (2016).
37. Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G. K. & Josic, K. The mechanics of state-dependent neural correlations. *Nat. Neurosci.* **19**, 383–393 (2016).
38. Churchland, M. M. et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369–378 (2010).
39. Wagner, M. J. et al. Shared cortex-cerebellum dynamics in the execution and learning of a motor task. *Cell* **177**, 669–682 (2019).
40. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).
41. Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
42. Keller, A. J., Roth, M. M. & Scanziani, M. Feedback generates a second receptive field in neurons of the visual cortex. *Nature* **582**, 545–549 (2020).
43. Bondy, A. G., Haefner, R. M. & Cumming, B. G. Feedback determines the structure of correlated variability in primary visual cortex. *Nat. Neurosci.* **21**, 598–606 (2018).
44. Zipsper, K., Lamme, V. A. & Schiller, P. H. Contextual modulation in primary visual cortex. *J. Neurosci.* **16**, 7376–7389 (1996).
45. Mashour, G. A., Roelfsema, P., Changeux, J. P. & Dehaene, S. Conscious processing and the global neuronal workspace hypothesis. *Neuron* **105**, 776–798 (2020).
46. Cohen, M. X. & Ranganath, C. Reinforcement learning signals predict future decisions. *J. Neurosci.* **27**, 371–378 (2007).
47. Bassett, D. S. & Bullmore, E. Small-world brain networks. *Neuroscientist* **12**, 512–523 (2006).
48. Oh, S. W. et al. A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Mice

The Stanford University Administrative Panel on Laboratory Animal Care approved all procedures using animals. For imaging studies of layer 2/3 neocortical pyramidal neurons in live mice, we used 4 male and 2 female triple transgenic *GCaMP6f-tTA-dCre (Rasgrf2-2A-dCre; Camk2a-tTA; Ai93)* mice developed by the Allen Institute. Mice were 10–16 weeks old at the time of surgery. Because there was only one group of mice, randomization and blinding procedures for the assignment of mice to different groups were not applicable. We did not use formal statistical methods to predetermine the sample size but instead replicated our main findings in all six mice used.

Surgical procedures

To prepare mice for in vivo imaging sessions, we performed surgeries while mice were mounted in a stereotaxic frame under isoflurane anaesthesia (1.5–2% isoflurane in O₂). To reduce post-operative inflammation and pain, we administered a preoperative dose of carprofen (5 mg kg⁻¹; subcutaneous injection into the mouse's lower back), which we repeated once a day for 3 days after the surgery. We created a cranial window by removing a 5-mm-diameter skull flap (centred at AP -2.5, ML 2.7) over the right cortical area V1 and surrounding cortical tissue. We covered the exposed cortical surface with a 5-mm-diameter glass coverslip (no. 1 thickness, 64-0700, CS-5R, Warner Instruments) that was attached within a circular steel annulus (1 mm thick, 5 mm outer diameter, 4.5 mm inner diameter, 50415K22, McMaster) and secured to the cranium using ultraviolet-light curable cyanoacrylate glue (Loctite, 4305). Using dental acrylic, we cemented a metal head-plate to the skull for head-fixation during imaging. In vivo brain imaging studies commenced at least 7 days after surgery.

Retinotopic mapping

To locate the boundaries of the visual cortical areas, we performed retinotopic mapping of the visual cortex in awake mice using widefield Ca²⁺ imaging by adopting a protocol that was used previously for retinotopic mapping by intrinsic signal imaging^{49–52}. As in all subsequent imaging experiments, we held mice on top of a 11.4-mm-diameter Styrofoam ball (Plasteel) using a two-point head-holder positioned under the objective lens of our custom-built epi-fluorescence microscope (see the 'Fluorescence microscope' section below; Fig. 1a). The Styrofoam ball floated on a thin layer of water within a plastic bowl of nearly identical diameter (Crittter-Cages), as previously described⁵³.

Mice viewed a visual stimulus comprising a drifting bar (10° wide) displayed on a video monitor positioned 13 cm from the left eye. The bar swept across the entire monitor in 14 s at a speed of 7° s⁻¹ and was filled internally with a contrast-reversing checkerboard pattern (0.035 per degree spatial frequency; 1.25 Hz temporal frequency of checkerboard reversal). The bar drifted either left, right, up or down on the monitor; each mouse viewed 100 repetitions of this stimulus for each direction of motion. The monitor remained grey for a 2 s interval between successive stimulus repetitions^{49,51}. Throughout the mapping session, we imaged baseline and evoked neocortical Ca²⁺ activity using the fluorescence microscope.

The visual stimulus used for mapping generally evoked retinotopic neural Ca²⁺ activity across the visual cortex, followed by a strong decline in Ca²⁺ activity below baseline levels. For each direction of stimulus motion, we computed the trial-averaged video of evoked Ca²⁺ activity, M (a three-dimensional matrix with spatial indices i and j , and a temporal index t), across all 100 stimulus repetitions, temporally aligned to the moment of stimulus onset. To map positions of the moving bar within the visual field to the corresponding anatomic coordinates within the visual cortical retinotopic maps, we calculated the phase of Ca²⁺ excitation within the i, j th pixel at each time t by approximating M with

a factorized model of a moving wave for each stimulus direction, so as to minimize the reconstruction error:

$$\text{Minimize}_{A,f,p} \sum_{i,j,t} (M_{ijt} - A_{ij} f(t - p_{ij}))^2.$$

Through this factorization, we approximated the average video M using a single waveform, f , with amplitude, A_{ij} , and phase, p_{ij} , at the i, j th pixel. We determined the values for the matrices, A and p , and the function, f , by using gradient descent to minimize the squared reconstruction error, summed over all pixels and time bins. We spatially smoothed the resulting phase maps using a Gaussian low-pass filter ($\sigma = 40 \mu\text{m}$) (Extended Data Fig. 1e).

On the basis of the smoothed phase maps determined for the vertical and horizontal directions of stimulus motion, we located the boundaries between V1 and the secondary visual areas (the medial visual (MV) and lateral visual (LV) cortical areas)⁴⁹. We inferred the locations of other cortical areas by aligning the Allen Brain Atlas cortical map⁵⁴ to the V1 boundaries determined in each mouse. Throughout the paper, for simplicity, we refer to the union of the lateromedial (LM) and anterolateral (AL) cortical areas as the LV, to the union of the anteromedial (AM) and posteromedial (PM) areas as the MV areas, and to the union of the rostralateral (RL) and anterior (A) areas as posterior parietal cortex (PPC). This grouping of the smaller secondary visual areas reduced to 8 the number of areas used in our subsequent analyses.

Training procedure and behaviour

We trained mice to perform the Go/No-Go task through successive stages of training (detailed below) that enabled us to gradually increase the complexity of the task performed by the mice while also ensuring that the association between visual stimuli and rewards remained stable. All of the mice in this study associated a Go stimulus with a horizontal grating orientation. To prevent light from the visual stimuli from entering the fluorescence collection pathway of the microscope, the stimuli used only the blue component of the RGB colour model, which was blocked by the fluorescence emission filter. We also placed a colour filter (Rosco, 382 Congo Blue) on the monitor screen. The mean luminance from the stimulus at the mouse eye was approximately 5×10^{10} photons mm⁻² s⁻¹, which is more than two orders of magnitude higher than the transition threshold to photopic vision in mice.

In the first stage, we trained water-deprived mice (target weight, 80% of initial body weight) to respond to a 100% contrast single drifting grating stimulus (2 s in duration; 2 Hz temporal frequency; 0.04 per degree spatial frequency; located within a 40° wide circle at the centre of a video monitor positioned 13 cm from the eye throughout all stages). In the first stage, mice learned that, by licking a spout during presentation of the Go stimulus, they would immediately receive a drop of 5% sucrose in water (~5 μl per drop). After a few days of training, mice that consistently licked only during Go trials progressed to the next stage of training.

In the second training stage, in addition to the Go stimulus, mice also viewed an orthogonal drifting grating stimulus or No-Go stimulus. Similarly to the first stage, mice were trained to respond on Go trials during the grating presentation, but we also included a grace period (1 s) at the onset of the grating stimuli that did not count towards a response. This allowed for some level of compulsive licking. After the grace period, if mice responded during No-Go stimuli, they received two aversive stimuli: (1) a small air puff (100 ms long) delivered to one eye of the mouse (contralateral eye to the stimulus); (2) simultaneously with the delivery of the air puff, the trial aborted and an 8 s timeout period occurred, during which the video monitor was held entirely grey at its mean luminance value. During this timeout, any additional lick(s) by the mouse resulted in the delivery of additional air puff(s). Once mice learned to perform the visual discrimination correctly on >75% of trials

by licking in response to the Go stimulus and not licking in response to the No-Go stimulus, training progressed to its next stage.

In the third training stage, we sought to create a separate response window so that rewards would not be provided at the same time as presentation of the visual stimuli. In this stage, mice learned to withhold their licks during stimulus presentation and to wait for a response period that was cued by an auditory tone (3.4 kHz; 100 ms duration). As in the second training stage, if mice licked during the visual stimulus they automatically received an air puff and a timeout (timeout duration was 3 s in the third training stage). As this training stage was the most challenging for the mice, we gradually increased the duration of the delay period either from session to session, or in three sub-blocks within one session, such that each mouse eventually performed the task with a delay of 0.5 s between the stimulus period (2 s duration) and the response period (3 s duration).

On a final day of training, we decreased the contrast of the moving gratings on both the Go and No-Go trials to between 50% and 12% to increase the proportion of error trials. Mice received only a single day of training on which the visual discrimination task was presented with this reduced level of visual contrast. By the end of training, all of the mice used for neural Ca^{2+} imaging studies performed the task with an accuracy of >75% with the low-contrast stimuli, for both Go and No-Go trials (Extended Data Fig. 1g, h; $83 \pm 3\%$ correct trials; mean \pm s.e.m.; $n = 6$ mice). Mice took 21–29 days of training (mean: 25 days; $n = 6$ mice) to reach the end of the training protocol.

Fluorescence microscope design

To image neural Ca^{2+} activity across 11 mouse cortical areas, we designed and built a custom wide-field fluorescence microscope with a field of view spanning 4 mm in diameter (Fig. 1a). For epifluorescence illumination, we used a light-emitting diode (LED) (Thorlabs, M470L2) with an emission spectrum centred in the 440–480 nm range. The imaging pathway comprised an objective lens (Leica, $\times 5.0$ Planapo 0.5 NA; 19 mm working distance; anti-reflection coated for 400–1,000 nm light; transmission >90% at 520 nm), a tube lens (75 mm focal length; Thorlabs, AC508-075-A-ML), a custom fluorescence filter cube (excitation filter, Semrock FF01-466/40-25; dichroic mirror, Semrock FF495-Di03, custom-sized to 35 mm \times 50 mm; emission filter, Semrock FF02-525/40, custom-sized to 30 mm \times 30 mm) and a scientific-grade CMOS camera (Hamamatsu ORCA-Flash4.0 V2 sCMOS). To control image acquisition, we used the HCLImage software (Hamamatsu), which communicated with the camera through an Active Silicon Firebird Camera Link Board.

To collect light from the LED, we used a 75 mm focal length focusing lens (Thorlabs LA1680, Thorlabs) to project convergent rays of excitation light at the back aperture of the microscope objective. We aligned the focusing lens to provide approximately uniform illumination across the field of view (5 mm diameter), that is, close to the regime of Kohler illumination, while also ensuring that the illumination rays were divergent as they entered the brain. The purpose of this illumination strategy was to create more intense illumination within neocortical layer 2/3 and to reduce fluorescence excitation within out-of-focus, deeper cortical layers. To improve the optical resolution at the periphery of the field of view, beyond the nominal 2-mm-diameter field of view of the objective lens, we reduced the effective numerical aperture by placing a 10-mm-diameter iris at the back aperture of the objective lens.

We built the opto-mechanical assembly using a combination of commercially available components (Thorlabs) and custom-designed mechanical parts machined in high-strength 7075 aluminium. The entire microscope was mounted on a manual vertical translation stage that enabled the user to conveniently adjust the image focus by moving the entire optical pathway of the microscope while the specimen was held immobile on the vibration-isolation table upon which the microscope was built.

Image acquisition and preprocessing

We acquired Ca^{2+} videos of neural activity (20 fps, $2,048 \times 2,048$ pixels) on the fluorescence microscope using 40–160 $\mu\text{W mm}^{-2}$ illumination. Custom software written in MATLAB (v.2013b) controlled the presentation of the visual stimuli to the mouse, ran the behavioural apparatus via a NI-USB 6008 card and triggered the start of video capture on the fluorescence microscope.

After video acquisition, we downsampled each video to $1,024 \times 1,024$ pixels and 10 fps. We next corrected videos for lateral movements of the brain using the Turboreg software package for image alignment⁵⁵. To remove scattered fluorescence and background fluorescence signals from neuropil or neural elements outside the focal plane, we applied a Gaussian spatial high-pass filter ($\sigma = 80 \mu\text{m}$) and calculated the video of relative fluorescence changes, $\Delta F(t)/F_0$, for each imaging session, where F_0 is the mean activity of each pixel over the entire session and $\Delta F(t)$ is the mean subtracted activity of each pixel at time t .

To quantify the slight lateral spatial displacements of the field of view between different imaging sessions, we computed the maximum projection image of each session's $\Delta F(t)/F_0$ video over its entire duration (around 1 h per session). We used the MATLAB *imregtform()* function to find the optimal 'similarity' transformations (translation, rotation and scaling) between the maximum projection image determined for the first imaging session and each of the other individual sessions. We aligned all Ca^{2+} videos to the video from the first session using this same set of transformations. Finally, we concatenated the aligned $\Delta F(t)/F_0$ videos from all sessions and proceeded to extract individual cells and their Ca^{2+} activity traces (see below; Extended Data Fig. 1).

Cell sorting

We extracted the activity of individual neurons from the concatenated $\Delta F(t)/F_0$ videos through the successive application of principal and independent analyses (PCA/ICA)⁵⁶. We divided the concatenated, preprocessed Ca^{2+} video from each mouse (about 1 TB in size) into 16 tiles; each tile comprised 256×256 pixels collectively covering about $1 \text{ mm} \times 1 \text{ mm}$ in the specimen plane. We ran PCA/ICA in parallel for all 16 tiles on 16 separate computing nodes (20 cores per node; 320 total cores; about 4 TB of random-access memory (RAM) for each video) and thereby identified Ca^{2+} activity traces and spatial filters for individual neurons. To isolate each cell soma, we thresholded each cell's spatial filter at 4 s.d. of its noise fluctuations (determined by fitting a Gaussian distribution to the negative values of each cell's spatial filter) and replaced all filter weights below this threshold with zeros. To attain a final set of Ca^{2+} -activity traces, we reapplied the truncated spatial filters to the $\Delta F(t)/F_0$ video (Extended Data Fig. 1).

To separate the sources of Ca^{2+} activity that represented individual cells from those that did not, for each mouse, we took 3 of the 16 image tiles and we manually identified individual neurons on the basis of both their morphologies and the temporal waveforms of their Ca^{2+} transients. To identify cells located within the other 13 tiles, we trained 3 different types of binary classifiers (support vector machine (SVM), generalized linear model (GLM) and neural network) to perform the classification by using the set of manually identified cells as training data and a set of 12 predefined cellular features that characterized a candidate neuron's morphology (spatial features: eccentricity, diameter, area, orientation, perimeter and solidity) and Ca^{2+} activity trace (mean peak amplitude of Ca^{2+} transients; signal-to-noise ratio between Ca^{2+} transients and baseline fluctuations; number of Ca^{2+} transients peaks that were 3 s.d. above baseline fluctuations; the number of Ca^{2+} transients peaks that were 1 s.d. above baseline fluctuations; the difference of the mean decay and mean rise times of the Ca^{2+} transients, normalized by the sum of these two values; and the FWHM of the average Ca^{2+} transient) to perform this classification. We used the trained classifiers to identify cells in the 13 remaining tiles based on a majority vote of the 3 classifier outputs. We manually checked that every cell

determined by this algorithm indeed met our visual inspection criteria to qualify as a neuron.

Event detection and definition of active cells

Using the fluorescence activity traces for the sources identified as neurons, we created binarized Ca^{2+} event traces for each cell (100 ms per time bin). To do this, we first subtracted the median level of fluorescence from each trace; we then calculated the s.d. of each cell's fluorescence fluctuations about baseline by fitting the statistical distribution of the activity trace's negative values to a Gaussian function constrained to have zero mean. To identify individual Ca^{2+} events, we looked for individual Ca^{2+} transients with peak amplitudes >4 s.d. above baseline fluctuations. The resulting binarized event traces had entries of '1' between the time at which the fluorescence amplitude of a Ca^{2+} transient surpassed 4 s.d. and the time at which the fluorescence amplitude started its decline back to baseline levels (Extended Data Fig. 1b). Entries were '0' for all other time bins. To account for slight day-to-day variations in the illumination, optical focal plane or amplitude of fluorescence fluctuations, we performed these computations separately for each imaging session.

To determine whether a cell was active during an individual imaging session, we counted the number of time bins in the session in which the cell's fluorescence emission was >3 s.d. above baseline fluctuations. We considered the cell to be 'active' if this number was >2 times greater than what would be predicted on the basis of a null hypothesis that the fluorescence variations simply reflected Gaussian-distributed noise (that is, the prediction that 0.27% of the time bins per session should have trace values >3 s.d. above baseline fluctuations) (Fig. 2a and Extended Data Fig. 1d).

Assessments of spatial alignment quality

To evaluate the quality of spatial registration between datasets from different imaging sessions, we computed the spatial cross-correlation functions between corresponding image patches, ($256 \mu\text{m} \times 256 \mu\text{m}$ in size) within the maximum projection images determined from the Ca^{2+} videos from the first imaging session and one of the subsequent sessions. We determined the slight day-to-day shifts in each patch's location by finding for each session the displacement value corresponding to the peak amplitude in the cross-correlation function (Extended Data Fig. 2a). By sliding the location of the $256 \mu\text{m} \times 256 \mu\text{m}$ patch used in this computation across the field of view, and computing the spatial cross-correlations for each location of the patch, we constructed maps of spatial displacement across the imaging field. These displacement maps revealed that our spatial alignments were almost perfect near the centre of the field of view (mean displacements, <1 pixel), and slightly deteriorated near the corners of the field of view (mean displacements ≈ 1 pixel).

To evaluate how these small imperfections in spatial registration might have affected alignments of cells and their identities across imaging sessions, we determined the displacement of each cell across sessions by examining $256 \mu\text{m} \times 256 \mu\text{m}$ image patches centred on each cell on each day of the experiment and then computing spatial cross-correlation functions as above. We determined each cell's day-to-day displacements in the datasets by identifying the maxima of these cross-correlations. This analysis showed that 98.5% of cells exhibit ≤ 1 pixel displacement across days (Extended Data Fig. 2b). We calculated each cell's mean displacement across all imaging sessions and plotted the cumulative distribution of cells' displacements by pooling the data from all of the mice (Extended Data Fig. 2c). For each cell, we also measured the distance to the nearest neighbouring cell and plotted the cumulative distribution of these values for all of the mice (Extended Data Fig. 2d). These two cumulative distributions revealed that only a small percentage ($\sim 2\%$) of nearest neighbor cell pairs were separated by $<5 \mu\text{m}$, whereas 95.4% of cells had mean day-to-day displacements that were $\leq 5\%$, indicating that slight imperfections in image alignment were highly unlikely to affect the registrations of cells' identities across days.

Analyses of single-cell coding

To characterize the extent to which individual neurons responded differentially to the two visual stimuli, we calculated the fidelity, d' , with which the two stimuli could be distinguished on the basis of a cell's stimulus-evoked dynamics:

$$d' = \frac{M_{\text{Go}} - M_{\text{No-Go}}}{\sqrt{0.5(\sigma_{\text{Go}}^2 + \sigma_{\text{No-Go}}^2)}},$$

where M_{Go} and $M_{\text{No-Go}}$ are mean values and σ_{Go}^2 and $\sigma_{\text{No-Go}}^2$ are variances of the cell's evoked Ca^{2+} dynamics (based on the binarized Ca^{2+} event traces) in response to Go and No-Go stimuli. We computed these quantities as trial-averages across either the stimulus, delay or response periods of the correctly performed trials, as specified in the figure captions. To allow even-handed comparisons between single-cell and neural population coding properties, for analyses of single-cell stimulus-evoked responses, we used the same time interval within the stimulus presentation period, [0.5 s, 2 s] after stimulus onset, that we used to train consensus decoders (see below). We also computed a distribution of d' values for a set of trial-shuffled datasets, denoted d'_{shuffle} . We created the set of trial-shuffled datasets by performing 1,000 random permutations of the Go and No-Go trial labels. We determined that an individual neuron coded significantly for stimulus identity during the stimulus, delay or response periods if the cell's d' value for that period was significantly greater than its d'_{shuffle} values for the same interval ($P < 0.01$; permutation test; $n = 710$ – $1,340$ trials). All analyses of single-cell coding, as well as those of neural ensemble coding and CCA modes were performed using only those trials on which the mouse's locomotor speed remained $<1 \text{ cm s}^{-1}$ throughout the trial.

Decoding neural population activity with optimal linear Fisher decoders

To quantify the information conveyed by neural ensemble dynamics about either the visual stimulus or the mouse's response, we used partial least squares analysis (PLS) as a supervised method for performing a dimensionality reduction, followed by optimal linear decoding in the space of reduced dimensionality, to determine d' , the fidelity with which the two stimuli or two responses could be distinguished on the basis of the activity patterns of the neural ensemble. The quantity (d')² is a discrete analogue of the Fisher information conveyed by the neural ensemble about the binary classification⁵⁷. Recent theoretical and computational work has shown that this approach for determining (d')² can yield accurate estimates even in the regime in which the number of experimental trials is far less than the number of neurons³.

For all decoding studies, we started by dividing all trials performed by each mouse into two distinct subsets, one used for decoder training and the other for decoder testing, and we represented the neural ensemble activity data in each subset using a three-dimensional tensor. The tensor elements, T_{ijk} , denoted the binarized activity of cell i on trial j at time bin k (Extended Data Fig. 3c). To train decoders, we used two different ways to convert these tensors into two-dimensional matrices.

In the first approach, we fixed the value of k in the tensor and trained a separate decoder based on the two-dimensional data matrix, X_{ij} , created for each time bin, k . We termed these decoders 'instantaneous decoders', because they enabled us to study the time-dependent dynamics of neural ensemble representations (Fig. 3a,b and Extended Data Fig. 3f,g). Notably, however, the instantaneous decoders of stimulus identity were largely stationary across the interval [0.5 s, 2 s] after stimulus onset. On the basis of this finding, we also pursued a second decoding approach that involved what we termed a single 'consensus decoder', which was designed to capture the non-dynamical aspects of the neural ensemble stimulus representations across all time bins in the [0.5 s, 2 s] interval.

In this second approach involving the consensus decoder, we took all 15 time bins of 100 ms each within the [0.5 s, 2 s] interval and

concatenated the data from these time bins along the trial index dimension, yielding a two-dimensional data matrix, X_{ij} . This matrix contained the data from the same number of cells as used for instantaneous decoding, but the effective number of trials was 15 times larger (Fig. 3c–j and Extended Data Fig. 3g). We used these matrices X_{ij} to train the consensus decoders of either stimulus identity or the mouse’s response.

An important consideration when training optimal linear Fisher decoders of either the instantaneous or consensus type was the fact that Fisher decoders require an estimate of the inverse of the noise covariance matrix of the neural ensemble activity patterns. When the number of recorded neurons surpasses the number of experimental trials, one cannot accurately estimate the individual elements of the noise covariance matrix. However, the principal eigenmodes and eigenvalues of this matrix can be determined accurately with a much smaller number of trials than neurons, which in turn enables accurate decoding and estimation of $(d')^2$ values³.

To achieve these estimates, as in our previous work, we first used PLS analysis to perform a supervised linear dimensionality reduction³ by identifying dimensions of the neural population activity in which the amplitude is correlated with the outcome of the binary classification task^{58,59}. The decoding strategy involved retaining a moderate number of these activity dimensions—while discarding the others—and then computing the optimal linear Fisher decoder and its associated d' value in this space of reduced dimensionality.

To train the optimal linear Fisher decoder for one of the binary classifications (that is, of either the stimulus identity or the mouse’s response), we split the two-dimensional data matrix, X_{ij} , as determined above, into two subsets, X^A and X^B , corresponding to the pair of conditions to be decoded. Specifically, the conditions A and B referred either to the two different visual stimuli or the two different possible responses by the mouse. Each row of the matrices X^A and X^B represented the neural activity data on a trial of type A or B , and each column represented the activity data from an individual neuron across all trials of this type. We randomly subsampled (with no replacement) the rows of X^A and X^B to create three distinct equally sized smaller data matrices, denoted X_{dr} , X_{tr} , and X_{te} , which we used for dimensionality reduction, decoder training and decoder testing, respectively, such that all of the data from any given trial were used in only one of these three matrices. Specifically, we used X_{dr} to find the set of PLS basis vectors, which comprised the columns of a coordinate transformation matrix, U . We transformed the training and testing datasets into the coordinate system defined by these PLS basis vectors:

$$\hat{X}_{tr} = X_{tr}U$$

$$\hat{X}_{te} = X_{te}U,$$

We systematically varied from 1–50 the number of PLS dimensions retained for the decoding analysis; the $\hat{\cdot}$ symbol indicates the vector space of reduced dimensionality. To determine the number of retained dimensions that yielded the highest decoding performance, we evaluated and optimized decoder performances through a cross-validation procedure (Extended Data Fig. 3c). Specifically, in the space of reduced dimensionality, we computed the optimal linear Fisher decoder, w_{opt} , from the training datasets, using the formula

$$w_{opt} = \hat{\Sigma}^{-1} \Delta \hat{\mu}, \quad (1)$$

where $\Sigma = 0.5\Sigma^A + 0.5\Sigma^B$ is the average noise covariance matrix and $\Delta \mu = \mu^A - \mu^B$ is the vector difference between the trial-averaged responses under conditions A and B . $\Delta \mu$ is also termed the ‘diagonal decoder’, namely a linear decoder that accounts for the mean responses under conditions A and B but not the covariances in these responses. We determined the binary decision boundary for the optimal linear decoder as the hyperplane normal to w_{opt} that bisected $\Delta \mu$. To attain

a decoder output or ‘score’ for an individual trial in the experiment, we projected the neural population dynamics from that trial onto w_{opt} and then subtracted $0.5\mu_{avg} \cdot w_{opt}$, where μ_{avg} is the mean of μ^A and μ^B , so the decoder score would have zero mean when averaged across a set of trials with equal numbers of A and B trials. We determined the binary classification using the sign of the score. Using the testing dataset, we estimated the discriminability of the two trial types, $(d'_{opt})^2$:

$$(d'_{opt})^2 = \frac{(w_{opt}^T \Delta \hat{\mu}_{te})^2}{w_{opt}^T \hat{\Sigma}_{te} w_{opt}}. \quad (2)$$

We repeated this process 100 times using 100 different random subsamplings of the trials for the construction of the dimensionality reduction dataset, the training dataset and the testing dataset.

To examine the extent to which visual stimulus encoding remained stationary over the course of the experiment, we trained an optimal ‘common decoder’ on the data recorded across all imaging sessions. To create the common decoder, we pooled all the data from each mouse and divided this aggregate set of data as described above into three subsets to be used for dimensionality reduction, decoder training and decoder testing. Given this division and using the procedures described above, we trained a consensus decoder for the interval [0.5 s, 2 s] after stimulus onset, yielding an across-day common decoder. We additionally assessed the values of $(d')^2$ for this common decoder on the testing datasets from the individual imaging sessions. This analysis revealed that the performance of the common decoder generally slightly surpassed that of decoders trained and tested on data exclusively from one imaging session (Fig. 3c and Extended Data Fig. 3i).

Analysis of error trials to distinguish neural coding of visual stimuli and mouse responses

On trials on which mice performed the Go/No-Go task correctly, the visual stimulus and the mouse’s response are perfectly correlated, precluding determinations of whether neural activity during the stimulus presentation is primarily evoked by the stimulus or also influenced by the mouse’s visual decision or information processing related to its upcoming response. To address this issue, we analysed error trials and trained decoders of neural ensemble activity that were sensitive to only the stimulus or only the animal’s decision, while keeping the other factor fixed.

For example, on Go trials the mouse could either lick (Hit) or not lick (Miss) (Fig. 1b). By training a ‘response decoder’ to discriminate between Hit and Miss trials on the basis of the neural activity during the stimulus presentation period, we estimated the encoded information about the mouse’s upcoming response while it observed the Go stimulus. As Hit trials were far more common than Miss trials, we randomly subsampled the set of Hit trials to construct unbiased datasets with equal numbers of Miss and Hit trials. Using these datasets, we trained consensus common decoders of neural population activity following the procedures discussed in the section above, as there were insufficient numbers of incorrectly performed trials to accurately train instantaneous decoders. Analyses of the visual stimulus period were based on the same interval, [0.5 s, 2 s] after stimulus onset, that was used to construct trial-type decoders. As the timing of the mouse’s responses differed from trial to trial and across trial types, we sought to retain sensitivity to the time dependence of coding by evaluating the response decoders’ $(d')^2$ values across the individual time bins of the trial structure. To construct the plots of Extended Data Figs. 3k and 4b–g, we identified the time bin of each trial with the maximum $(d')^2$ value and used that $(d')^2$ value when tabulating the results across trials and mice. Our decoding results revealed distinct patterns of neural activity during Go stimulus presentations that were predictive of the mouse’s upcoming response. We also executed an identical decoding analysis using equally sized datasets constructed from the neural activity recorded on No-Go trials (that is, Correct Rejection and False Alarm

trials). However, in this case, we did not find neural activity patterns during stimulus presentation that predicted the mouse's response (Extended Data Figs. 3k and 4e). As the response decoders trained on Go and No-Go trials were constructed using equally sized datasets, the differences in their performances cannot be readily explained as due to a discrepancy in statistical power.

To determine whether visual stimulus coding during stimulus presentation might have been affected by the mouse's upcoming response, we trained and evaluated separate common consensus stimulus decoders for Lick trials (False Alarm and Hit) and No-Lick trials (Correct Rejection and Miss), using the same methods as for response decoders and with equally sized datasets that were constructed through subsampling. This analysis yielded no evidence that the quality of stimulus representations was impacted by the mouse's upcoming response (Extended Data Figs. 3k and 4b).

Calculations of information redundancy across cortical areas

To assess the extent to which Fisher information about the stimulus was represented independently across different cortical areas, we examined inter-area correlations in the output scores of the instantaneous neural activity decoders (see above). We quantified these correlations separately for the two types of correctly performed trials and then averaged the resulting correlation coefficients.

The results revealed that fluctuations in neural ensemble activity along the stimulus coding direction were strongly correlated between the different sensory areas just after stimulus onset and then progressively decayed (Fig. 4a–c and Extended Data Fig. 6). If information were represented independently in the different cortical areas, the sum of the information encoded in each of the individual brain areas would equal that encoded in the aggregate of all the brain areas²⁵. Positive correlations in the decoder scores from different brain areas can reflect redundancy (Fig. 4d) such that this equality is not met and there are shared copies of the same information²⁵:

$$\text{Redundancy} = \frac{\sum_{\text{areas}} (d')^2_{\text{area}}}{(d')^2_{\text{all areas}}} \quad (3)$$

Determination of noise correlations among neuron pairs

To measure noise correlations between pairs of similarly tuned neurons, we trained instantaneous population decoders of the stimulus based on the neural activity recorded in each mouse on all trials performed correctly (see above). We selected cells that significantly contributed to each decoder by identifying those cells with decoder weights that deviated >2 s.d. from the mean value across the entire set of cells considered (Fig. 3g–j). We divided the resulting set of cells into two groups based on the sign of the individual cells' mean-subtracted decoder weights as an indicator of similarity in the cells' tuning to the visual stimulus. We then computed the noise correlation coefficients characterizing the joint activity fluctuations of pairs of cells around their mean responses. We averaged the values of these coefficients over the two types of correctly performed trials. The time dependence of these correlations closely resembled that of the noise correlations in decoder scores across brain areas (see above).

In our analysis, we did not find substantial noise correlations between cells with dissimilar stimulus tuning or between cells without stimulus tuning. This is in agreement with our past findings in untrained mice viewing moving grating stimuli that differed by 60° in orientation³ but, here, with trained mice actively performing a task involving an orthogonal pair of moving grating stimuli, the differences between the distributions of noise correlation coefficients between cell pairs with similar and dissimilar stimulus tuning were more substantial^{43,60} (Fig. 3m).

To estimate the time-dependent mean variability, $\sigma^2(t)$, of individual neuronal responses in each mouse, we computed the variance in the activity level of each cell at time, t , relative to stimulus onset,

across the set of all correctly performed Go and No-Go trials. We averaged the results across all cells and both trial types. To compute the time-dependent Fano factor across the set of all neurons (Extended Data Fig. 5e), we divided $\sigma^2(t)$ by $\mu(t)$, the cells' mean response at time t , averaged over all cells and correctly performed trials. Both $\sigma^2(t)$ and the Fano factor declined after stimulus onset, consistent with previous studies³⁸ (Extended Data Fig. 5e).

Determinations of information saturation in large neural ensembles

Previous theoretical and recent experimental work has shown that the Fisher information encoded in the dynamics of a cortical neural ensemble saturates at large ensemble sizes, due to the existence of eigenvectors of the noise covariance matrix with eigenvalues that grow linearly in the limit of large ensemble size^{3,5,14,25} (Extended Data Fig. 5a). To characterize this information saturation at each time bin after stimulus onset, we trained instantaneous decoders of the visual stimulus based on the activity of a subset of the neurons recorded in each brain area. We systematically varied the size of this subset and measured the encoded information using the decoder $(d')^2$ values for each ensemble size, as averaged over 100 random selections of neurons for each time bin during which the entire cell population significantly encoded information about the stimulus ($P < 0.01$; permutation test; $n = 710\text{--}1,340$ trials). We normalized the $(d')^2$ values from each time bin to the total information encoded by all neurons during this same time bin.

In agreement with recent studies of V1^{3,25}, in all of the cortical areas examined here the information encoded by a cell ensemble saturated at large ensemble sizes (Extended Data Fig. 5a). Furthermore, just after stimulus onset, this saturation occurred at much smaller neural ensembles compared with later on in the trial. As stimulus presentation proceeded, the functional dependence of $(d')^2$ on ensemble size became more similar to the form observed in trial-shuffled datasets (Fig. 3k and Extended Data Fig. 5b, c).

To estimate the sensitivity of the ensemble neural code to the hypothetical loss of one neuron, we determined the number of neurons of which the loss would result in a 10% decrement in the total information encoded by the cell population. We rescaled the result to express the information loss per cell removed (Extended Data Fig. 5h).

Determinations of the similarity between pairs of vector subspaces

To assess the similarity between two K -dimensional subspaces (Extended Data Figs. 3e and 5j), we first calculated the $K \times K$ matrix $S = U^T V$, where U and V are $N \times K$ matrices of which K orthonormal columns form a basis for each subspace. We then performed a singular value decomposition of S and determined the subspace similarity as the mean of the K singular values. This calculation yields zero for orthogonal subspaces and one for identical subspaces. As each singular value is the cosine of a canonical angle between the two subspaces, this measure is equivalent to the mean of the cosines of the K canonical angles.

Assessments of how day-to-day drifts in neural encoding relate to trial-to-trial activity fluctuations

To assess how the day-to-day variations in stimulus-evoked neural responses related to the trial-to-trial variations in these responses within individual imaging sessions, we first rescaled each neuron's activity trace to have zero mean and unit variance on each day of the experiment. Using these traces, we calculated the noise covariance matrix of the stimulus-evoked neural responses on each day, and we averaged these matrices across the two trial types. To identify the principle directions of the trial-to-trial activity fluctuations on each day, we performed an eigenvector decomposition of each of the averaged covariance matrices.

To examine how the day-to-day variations in the neural representations related to the trial-to-trial activity fluctuations, we projected the

changes between successive days in the mean neural ensemble response on each trial type onto the eigenvectors of the noise covariance matrix for the first day in each pair of consecutive days (we obtained similar results if we alternatively chose the eigenvectors from the second day of each pair). We averaged the results over both stimuli and all pairs of consecutive days. As control, we performed the same analysis with trial-shuffled datasets, in which the noise covariance matrix was rendered isotropic by permuting the activity traces of each cell across trials of the same stimulus type. The results showed that day-to-day drifts in the neural ensemble representations of the stimuli were significantly aligned with the principal directions of the trial-to-trial variations within individual days (Fig. 3f and Extended Data Fig. 4a). We obtained similar results when we projected the day-to-day changes in the visual stimulus tuning curve onto the eigenvectors of the within-day, noise covariance matrix. A theoretical explanation for how this observation can enable optimal decoders to be robust across days, and an explanation of how this alignment between within-day fluctuations and across-day changes in mean neural ensemble responses can arise mechanistically in a simple network model without any fine tuning are provided in the Supplementary Information.

Effects of correlated noise in a two-layer feedforward network model of the visual cortex

To guide our examinations of how redundant information coding across different neural ensembles is related to correlated fluctuations in activity that reflect neuronal connectivity patterns (Fig. 3o), we analysed a two-layer feedforward network model, also discussed elsewhere³. This network comprises an input layer of ‘sensory neurons’ and an output layer of ‘cortical neurons’, of which the activity levels are denoted by the vectors \mathbf{r} and \mathbf{s} , respectively, and related by the expression

$$\mathbf{r} = \mathbf{F}(W\mathbf{s} + \boldsymbol{\xi}^{\text{in}}) + \boldsymbol{\xi}^{\text{out}}.$$

Here $\boldsymbol{\xi}^{\text{in}}$ and $\boldsymbol{\xi}^{\text{out}}$ are zero-mean Gaussian-distributed additive noise vectors that represent the stochastic components of the input and output activity levels, W denotes the connection matrix between the two layers, and \mathbf{F} is a nonlinear transfer function relating the net input and output levels of activity. We approximate the response to a specific stimulus A via a Taylor expansion:

$$\mathbf{r}^A = \mathbf{F}(W\mathbf{s}^A) + F'(W\mathbf{s}^A)^T \boldsymbol{\xi}^{\text{in}} + \boldsymbol{\xi}^{\text{out}},$$

where the prime symbol denotes the first derivative. As both $\boldsymbol{\xi}^{\text{in}}$ and $\boldsymbol{\xi}^{\text{out}}$ have zero means, the mean output response to this specific stimulus is $\boldsymbol{\mu}^A = \mathbf{F}(W\mathbf{s}^A)$ where \mathbf{s}^A is the mean activity evoked in the sensory layer by stimulus A . Under these assumptions, the noise covariance matrix between neurons in the cortical layer is:

$$\Sigma^A = G^A W \Sigma^{\text{in}} W^T G^A + \Sigma^{\text{out}},$$

where G^A is a diagonal matrix of which the elements denote the linear gain of each neuron around stimulus A , as determined from the function F' . If all neurons operate at similar gains (assumed to be 1 here for simplicity), and if the noise terms $\boldsymbol{\xi}^{\text{in}}$ and $\boldsymbol{\xi}^{\text{out}}$ are uncorrelated between neurons, independent of the stimulus, and have variances, σ_{in}^2 and σ_{out}^2 , that are uniform for all cells in each layer, then:

$$\Sigma = \sigma_{\text{in}}^2 W W^T + \sigma_{\text{out}}^2 I, \quad (4)$$

where I is the identity matrix. To compute the $(d')^2$ value for distinguishing between two distinct stimuli using an optimal linear decoder of activity in the output layer, the application of equation (1) above leads to:

$$(d')^2 = \Delta \boldsymbol{\mu}^T \Sigma^{-1} \Delta \boldsymbol{\mu} = \Delta \mathbf{s}^T W^T (\sigma_{\text{in}}^2 W W^T + \sigma_{\text{out}}^2 I)^{-1} W \Delta \mathbf{s}. \quad (5)$$

Our previous analysis of this model³ shows that, if we replace W in equation (5) by its singular value decomposition, the minimum number of neurons, $N_{0.5}^\alpha$, needed on average to extract >50% of the encoded information along each left-singular vector, \mathbf{u}_α , of W is determined by:

$$N_{0.5}^\alpha = \frac{1}{\bar{d}_\alpha^2} \frac{\sigma_{\text{out}}^2}{\sigma_{\text{in}}^2}, \quad (6)$$

where \bar{d}_α^2 is the square of the α th largest singular value of W , divided by the total number of cortical neurons. From equation (4) we can also estimate the average value of the diagonal ($\bar{\Sigma}_{ii}$) and non-diagonal ($\bar{\Sigma}_{ij}$) elements of the noise covariance matrix:

$$\bar{\Sigma}_{ii} = \sigma_{\text{in}}^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle + \sigma_{\text{out}}^2 \quad (7)$$

$$\bar{\Sigma}_{ij} = \sigma_{\text{in}}^2 \langle \mathbf{w}_i, \mathbf{w}_j \rangle \quad (8)$$

where $\langle \mathbf{w}_i, \mathbf{w}_i \rangle = \frac{1}{N} \sum_{i=1}^N \mathbf{w}_i^T \mathbf{w}_i$ is a mean amplification factor, averaged over the N singular vectors of W (where N is the number of cells in the output layer) and $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} \mathbf{w}_i^T \mathbf{w}_j$ is the mean similarity between the receptive fields of cells in the output layer.

Dividing equation (7) by equation (8) yields:

$$\frac{\sigma_{\text{out}}^2}{\sigma_{\text{in}}^2} = \langle \mathbf{w}_i, \mathbf{w}_j \rangle \frac{\bar{\Sigma}_{ii}}{\bar{\Sigma}_{ij}} - \langle \mathbf{w}_i, \mathbf{w}_i \rangle. \quad (9)$$

Finally, substituting equation (9) into equation (6) yields:

$$N_{0.5}^\alpha = \frac{\langle \mathbf{w}_i, \mathbf{w}_j \rangle \frac{\bar{\Sigma}_{ii}}{\bar{\Sigma}_{ij}} - \langle \mathbf{w}_i, \mathbf{w}_i \rangle}{\bar{d}_\alpha^2} \frac{1}{\bar{d}_\alpha^2}. \quad (10)$$

Equation (10) shows how the number of cells in the output layer needed to extract half-maximal information is related to the basic structure of the connectivity matrix, W (Fig. 3o).

Empirical analyses of redundancy and noise covariance in cortical ensembles

To study whether equation (10) held empirically in our datasets, we computed the ratio, $\bar{\Sigma}_{ii}/\bar{\Sigma}_{ij}$, from our recordings of cortical neurons and studied its relationship to $N_{0.5}$. In equation (10), $N_{0.5}^\alpha$ is related to an individual eigenvector of the connectivity matrix, W . The value of $N_{0.5}$ for an entire neural ensemble will be primarily determined by those eigenvectors of the connectivity matrix that make significant contributions to stimulus coding. As we do not have direct access to W , the connectivity matrix of the mammalian brain, to test equation (10) we estimated the noise properties of neurons that contributed significantly to stimulus coding.

To estimate Σ , we computed the noise covariance for each stimulus separately and then averaged the results for both stimuli (Go and No-Go). We estimated $N_{0.5}$ during the stimulus interval separately for each time bin (Fig. 3l; see above for detailed methods). In our experiment, the $N_{0.5}$ values and noise correlation coefficients varied over time during the stimulus presentation period. Equation (10) suggests that this time dependence should be constrained such that there is a linear relationship between $N_{0.5}$ and $(\bar{\Sigma}_{ii}/\bar{\Sigma}_{ij})$ at all time points. To test this, for each time bin, we plotted the empirically determined values of $N_{0.5}$ (Fig. 3l) against the ratio, $\bar{\Sigma}_{ii}/\bar{\Sigma}_{ij}$, computed across the set of all cells that significantly encoded the stimulus type (see above for how we identified these neurons). The results were strikingly consistent with the linear relationship predicted by equation (10) (Fig. 3o). The slope of the linear relationship was similar for all of the mice in the experiment, which presumably reflects conserved properties of the anatomical neural connectivity within the mouse visual pathways, such

as the degree of overlap in the receptive fields of nearby cells and the amplification factors across different stages of visual processing.

Analysis of canonical noise correlations

To examine the structure of correlated activity fluctuations across different cortical areas and their relationships to the representation of information, we used CCA⁶¹ to study the co-variations of activity fluctuations within pairs of brain areas. For each trial type, we computed the trial-by-trial fluctuations in stimulus-evoked activity by subtracting from each fluorescence Ca^{2+} trace the mean Ca^{2+} activity trace, averaged over all trials. We concatenated the traces representing these fluctuations across trials that the mouse performed correctly. For a given pair of brain areas, we represented the dynamics in the two areas with matrices, X and Y . These matrices were $N_t \times N_1$ and $N_t \times N_2$ in size, where N_t was the total number of time points after the concatenation, and N_1 and N_2 were the numbers of cells detected in each brain area. We standardized these zero-mean matrices of fluctuations X and Y by scaling each matrix column to have unit variance.

Following the standard approach in CCA, we identified two sets of loading vectors, $\{\mathbf{w}_i\}$ and $\{\mathbf{v}_i\}$, termed here as CCA modes, each of which was an activity mode within one of the two neural ensembles (that is, with N_1 and N_2 elements, respectively). The index $i \in \{1, 2, 3, \dots, \text{minimum}(N_1, N_2)\}$ denoted the individual modes, which we determined such that the projections of the neural activity fluctuations, X and Y , onto \mathbf{w}_i and \mathbf{v}_i , were maximally correlated between the two ensembles,

$$\text{Maximize}_{\mathbf{w}_i, \mathbf{v}_i} (X\mathbf{w}_i)^T (Y\mathbf{v}_i), \quad (11)$$

subject to the normalization constraint, $\mathbf{w}_i^T X^T X \mathbf{w}_i = \mathbf{v}_i^T Y^T Y \mathbf{v}_i = 1$. Given this normalization condition, the quantity $(X\mathbf{w}_i)^T \cdot (Y\mathbf{v}_i)$ equals the correlation coefficient of the activity modes, $Y\mathbf{v}_i$ and $X\mathbf{w}_i$, in the two different brain areas. After finding the first CCA mode ($i = 1$), we identified successive modes in an iterative manner. Specifically, for all previously identified CCA modes we removed the CCA fluctuations, $Y\mathbf{v}_i$ and $X\mathbf{w}_i$, respectively, from X and Y . We applied equation (11) to the residuals and thereby identified a set of orthonormal fluctuation modes with correlation coefficient values that progressively declined with the index, i . To identify the maxima specified by equation (11), we first randomly initialized the vectors \mathbf{w}_i and \mathbf{v}_i while constraining them to have unity length. We then found values of \mathbf{w}_i and \mathbf{v}_i that maximized the objective function in equation (11) by performing an alternating optimization⁶².

To create training and validation datasets, we randomly divided the full datasets into two subsets with equal numbers of trials, with all the data from each trial used only in one of the two subsets. We used the first subset to find the top 20 CCA modes for all pairs of cortical areas. We used the second subset of trials to determine the inter-area correlation coefficients of the fluctuations in each of the CCA modes; this revealed significant correlated fluctuations in the test dataset with no signs of overfitting (Extended Data Fig. 7d). We also performed a CCA of trial-shuffled datasets. By comparing the correlation coefficients for CCA fluctuations in the real data with those observed across 100 different trial-shuffled datasets, we determined that the correlation coefficients in the real data were significantly larger than expected by chance ($P < 0.01$; permutation test; $n = 710$ – $1,340$ trials; 525 cells per brain area on average, range: 31–2,297 cells; Extended Data Fig. 7a).

We also measured the amplitude of canonical correlations separately for Go and No-Go trials and found that, on average, the correlation coefficients had similar values for the two stimulus types (Extended Data Fig. 7d). Thus, for most of our analysis, to simplify visualization of the data, we combined the sets of mean-subtracted activity traces for the two stimuli and identified a single set of CCA modes between each pair of brain areas, independent of the stimulus type.

As a control analysis to ensure that the inter-area activity fluctuations that we had identified had not artifactually arisen from slight errors in

determining the boundaries between brain areas, we performed CCA analysis on a control dataset in which we excluded all cells located $< 60 \mu\text{m}$ to the other brain area under consideration. These exclusions did not notably modify the amplitudes of correlated fluctuations or other aspects of our findings (Extended Data Fig. 7e).

To assess how the CCA correlation coefficients varied as a function of time relative to stimulus onset, for each pair of brain areas we projected the neural activity at different time bins onto the CCA modes and computed the correlation coefficient using the validation dataset; this yielded different values of the correlation coefficients for each time bin (Extended Data Fig. 8a). Across most of the visual stimulation period, the CCA fluctuations exhibited significantly greater correlation coefficients in the real than in trial-shuffled datasets ($P < 0.01$, permutation test, $n = 710$ – $1,340$ trials 525 cells per brain area on average; range, 31–2297 cells).

To examine how the brain's fluctuation modes might change at the onset of visual stimulation, we first used CCA to identify a distinct set of CCA modes of the neural ensemble dynamics during ITIs, within the period $[-2 \text{ s}, 0 \text{ s}]$ relative to stimulus onset. We then compared these CCA modes to those found within the visual stimulus period $[0 \text{ s}, 2 \text{ s}]$. To do this, once we had identified CCA modes during visual stimulus presentation using training datasets, we extended the temporal range of the validation datasets to include the $[-0.5 \text{ s}, 0 \text{ s}]$ interval. Conversely, once we had identified CCA modes during the ITIs, we extended the temporal range of the validation datasets to include the $[0 \text{ s}, 0.5 \text{ s}]$ interval. We found that the correlation coefficient values of the ITI CCA modes declined after stimulus presentation, whereas those for the stimulus period CCA modes sharply increased shortly after stimulus onset (Extended Data Fig. 8a). For each CCA mode index, i , we also compared the directions of the mode vectors within the neural population activity vector space for the two different sets of CCA results by determining the cosines of the angles between the i th CCA mode vectors from before versus after visual stimulus onset (Extended Data Fig. 8b).

For comparison, we trained CCA modes using the data from the entire $[-2 \text{ s}, 2 \text{ s}]$ interval, subsampled so that the training datasets were equally sized to those used to train the ITI and stimulus CCA modes from the $[-2 \text{ s}, 0 \text{ s}]$ and $[0 \text{ s}, 2 \text{ s}]$ intervals, respectively. At stimulus onset, many of these CCA modes exhibited either a rise or a decline in their canonical correlation coefficients, consistent with the results obtained when we trained CCA modes separately for the $[-2 \text{ s}, 0 \text{ s}]$ and $[0 \text{ s}, 2 \text{ s}]$ intervals. However, the values of the canonical correlation coefficients for the modes trained for the $[-2 \text{ s}, 2 \text{ s}]$ interval were generally less than those of the CCA modes trained separately for the stimulus presentation and ITI presentations, suggesting that the implicit assumption in CCA of statistical stationarity does not hold at stimulus onset and that there is a bona fide transition in the noise correlation structure of cortical activity at stimulus onset.

Simulations of multi-area neural fluctuations

To study how neural connectivity can give rise to CCA modes that share information between brain areas, we modelled the linear network schematized in Extended Data Fig. 9f with $N_c = 500$ cells in each of one 'early visual area' and three 'cortical areas' (termed A , B and C). Neural activity in the early visual area, E , were set by

$$\mathbf{E} = \nu \mathbf{S} + W_{DE} (u\mathbf{M}) + \boldsymbol{\xi}^E,$$

where \mathbf{S} and \mathbf{M} were 500-dimensional unit vectors (with fixed values in each simulation) representing input patterns of neural ensemble activity encoding the stimulus and the mouse's response, respectively, and ν and u were binary variables with values of either -1 or 1 that represented the two stimulus and response conditions. W_{DE} was a linear low-rank projection matrix from the space of the decision variable to that of the neural activity levels; we systematically varied the rank, k ,

of this matrix from 1–10 across multiple runs of the simulation. Specifically, W_{DE} was the outer product of two $N_c \times k$ matrices in which all of the elements were randomly and independently chosen from a zero-mean unit variance Gaussian distribution, and each column of these two matrices was normalized to have an L2-norm of 1. ξ^E was an additive noise vector in which the individual elements were independently drawn from identical zero-mean Gaussian distributions with variance = $1/N_c$. The neural dynamics in areas A , B and C differed in that, instead of directly receiving stimulus information, they received it indirectly via a low-rank linear projection from area E . For example, activity levels in area A were set by

$$\mathbf{A} = W_{EA} \mathbf{E} + W_{DA} (u\mathbf{M}) + \xi^A,$$

where W_{EA} and W_{DA} are linear low-rank projection matrices; analogous equations governed the dynamics for areas B and C . As with ξ^E , the elements of the additive noise terms, ξ^A , ξ^B and ξ^C were independently drawn from identical zero-mean Gaussian distributions with variance = $1/N_c$. We systematically varied the ranks of the matrices W_{DE} , W_{EA} , W_{DA} , W_{EB} , W_{DB} , W_{EC} and W_{DC} to have values between 1 and 10; for each of the 10 different values of k , we repeated the simulations 25 times with different sets of randomly chosen matrix elements and different randomly chosen values for \mathbf{S} and \mathbf{M} . We simulated each of the 250 models for 20,000 trials; on each trial, we chose the stimulus and decision variables, u and v , randomly and independently of each other. We used the methods described above to find the CCA modes of each model (Extended Data Fig. 9g–i).

Simulations of small-world networks

As shown in Extended Data Fig. 9f, g, the global transmission of a common decision signal to multiple cortical areas can produce a global CCA mode that is shared among all pairs of cortical areas, similar to what we found in the real neural recordings. To examine whether a global CCA mode can also arise in the absence of a globally transmitted signal, we modelled networks with 11 brain areas that were interconnected according to a small-world connectivity rule⁶³, with unidirectional connections^{30,64,65} (Extended Data Fig. 9b).

We simulated 30 different networks with varying degrees of interconnectivity and varying levels of randomness and regularity in the pattern of connections. For each network, we set the graph of connections by arranging the 11 brain areas in a ring formation. We then created unidirectional projections to each brain area from its K nearest neighbours on the ring (that is, from $K/2$ neighbouring areas on both sides of each brain area). To introduce randomness into the connectivity pattern, the brain areas sending each of these unidirectional projections were then randomly re-assigned with probability, P , to a different brain area that was randomly selected with uniform probability $1/(11 - K)$ from among those areas that had originally lacked such a projection.

Within each area, there were 500 neurons, of which the activity levels were a linear function of the neural activity in the brain areas from which they received inputs:

$$\mathbf{X}^n(t) = \alpha \xi^n(t) + \frac{\beta}{\sqrt{K}} \sum_m^{N_{\text{area}}} a_{m,n} W^{m,n} \mathbf{X}^m(t-1).$$

Here $\mathbf{X}^n(t)$ is a vector of 500 elements that represent the activity of the 500 cells in the n th brain area at time t . $\xi^n(t)$ is an additive noise term for the n th area, in which the individual elements at time t were independently drawn from identical zero-mean Gaussian distributions with a variance of 4×10^{-4} . $W^{m,n}$ is a 500-rank projection matrix from area m to area n , in which all the elements were chosen randomly and independently from a zero-mean unit variance Gaussian distribution; all of the columns of $W^{m,n}$ were normalized to have an L2 norm of 1. $a_{m,n} = 1$ if and only if there was an edge from node m to node n in the small-world graph; otherwise $a_{m,n} = 0$. The parameters α and β were gain

factors; their relative amplitudes determined the degree of coupling between areas.

In general, $\beta < 1$, because increasing the value of β too close to 1 can cause the whole network to enter a global oscillation mode with a period of 2 cycles. With further increases of $\beta \rightarrow 1$, the network becomes unstable. We therefore selected β so as to provide strong coupling between brain areas while avoiding the fast global oscillatory mode. We simulated this linear system for all possible combinations of $K \in \{2, 4, 6, 8, 10\}$ and $P \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. To reproduce CCA modes with similar correlation coefficients to those we had observed in the real cortical recordings, we set $\alpha = 0.01$ and $\beta = 0.9$. For each set of K and P values, we initialized the neural activity levels, $\mathbf{X}^n(t)$, in the model with zero-mean Gaussian noise with variance = 4×10^{-8} and ran the simulation for 50,000 time points. To avoid effects arising from initial transients, we omitted from all analyses the data from the first 500 time steps.

Data and statistical analyses

We performed all data and statistical analyses using MATLAB (version R2019a; MathWorks). All statistical tests were two-sided, except for permutation tests, which were one-sided. All signed-rank tests were Wilcoxon signed-rank tests.

Computational simulations

We performed all simulations using MATLAB (version R2019a; MathWorks).

Reporting summary

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

Data availability

The data supporting the findings of this study are available from the corresponding authors on reasonable request.

Code availability

We used open source software routines for image registration⁵⁵ (<http://bigwww.epfl.ch/thevenaz/turboreg/>) and PLS analysis (<https://www.mathworks.com/matlabcentral/fileexchange/18760-partial-least-squares-and-discriminant-analysis>). Software code for extracting individual neurons and their calcium activity traces from calcium videos by using principal component and then independent component analyses⁵⁶ is available online for free (<https://www.mathworks.com/matlabcentral/fileexchange/25405-emukamel-cellsort>), although, for convenience, we used a commercial version of these routines (Mosaic software, v.0.99.17, Inscopix). We used MATLAB (v.2019a) to write all other analytical routines. The primary software code used to support the findings of the study is available at Zenodo (<https://doi.org/10.5281/zenodo.6314932>).

49. Garrett, M. E., Nauhaus, I., Marshel, J. H. & Callaway, E. M. Topography and areal organization of mouse visual cortex. *J. Neurosci.* **34**, 12587–12600 (2014).
50. Kalatsky, V. A. & Stryker, M. P. New paradigm for optical imaging: temporally encoded maps of intrinsic signal. *Neuron* **38**, 529–545 (2003).
51. Marshel, J. H., Garrett, M. E., Nauhaus, I. & Callaway, E. M. Functional specialization of seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
52. Zhuang, J. et al. An extended retinotopic map of mouse cortex. *eLife* **6**, e18372 (2017).
53. Lecoq, J. et al. Visualizing mammalian brain area interactions by dual-axis two-photon calcium imaging. *Nat. Neurosci.* **17**, 1825–1829 (2014).
54. Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
55. Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).
56. Mukamel, E. A., Nimmerjahn, A. & Schnitzer, M. J. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* **63**, 747–760 (2009).
57. Kanitscheider, I., Coen-Cagli, R., Kohn, A. & Pouget, A. Measuring Fisher information accurately in correlated neural populations. *PLoS Comput. Biol.* **11**, e1004218 (2015).
58. Barker, M. & Rayens, W. Partial least squares for discrimination. *J. Chemometr.* **17**, 166–173 (2003).
59. Wold, H. in *Multivariate Analysis* (ed. Krishnajah, P. R.) 391–420 (Academic, 1966).

Article

60. Kohn, A. & Smith, M. A. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* **25**, 3661–3673 (2005).
61. Hotelling, H. in *Breakthroughs in Statistics* Vol. 2 (eds S. Kotz & N.L. Johnson) 162–190 (Springer, 1992).
62. Witten, D. M. & Tibshirani, R. J. Extensions of sparse canonical correlation analysis with applications to genomic data. *Stat. Appl. Genet. Mol. Biol.* **8**, Article28 (2009).
63. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
64. Honey, C. J., Kotter, R., Breakspear, M. & Sporns, O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl Acad. Sci. USA* **104**, 10240–10245 (2007).
65. Lu, J., Yu, X., Chen, G. & Cheng, D. Characterizing the synchronizability of small-world dynamical networks. *IEEE Trans. Circ. Syst.* **51**, 787–796 (2004).

Acknowledgements We thank B. Ahanonu, A. Christensen, H. Kim, T. Rogerson, A. Shai and A. Tsao for conversations, and H. Zeng for providing transgenic mice. We acknowledge research support from HHMI (to M.J.S.); the Stanford CNC Program (to M.J.S.); DARPA (to M.J.S.); NIH BRAIN Initiative grant 1UF1NS107610-01 (to M.J.S.); the NSF NeuroNex Program (to M.J.S.); an NSF CAREER Award (to S.G.); the Burroughs-Wellcome (to S.G.), McKnight (to S.G.), James S. McDonnell (to S.G.) and Simons (to S.G. and M.J.S.) foundations; and a Stanford Graduate Fellowship (to O.R.).

Author contributions S.E., J.A.L., S.G. and M.J.S. designed the study. S.E., J. Lecoq, O.R., T.T., C.I. and M.J.S. designed the experiments. J. Lecoq, O.R. and M.J.S. designed the microscope. J. Lecoq, C.I., S.E. and T.T. trained mice and acquired the data. S.E., S.G. and M.J.S. designed the analyses. S.E. analysed the data and performed the network simulations. S.G. did the theoretical and computational analyses of the Supplementary Information. Y.Z. and J. Lecoq. performed mouse surgeries. Y.Z. and J. Li performed histological analyses and husbandry. S.E., S.G. and M.J.S. wrote the paper. All of the authors edited the paper. S.G. and M.J.S supervised the research.

Competing interests M.J.S. is a scientific co-founder of Inscopix, which produces the Mosaic software used to identify individual neurons in the Ca²⁺ videos. J. Lecoq. is also an Inscopix stockholder.

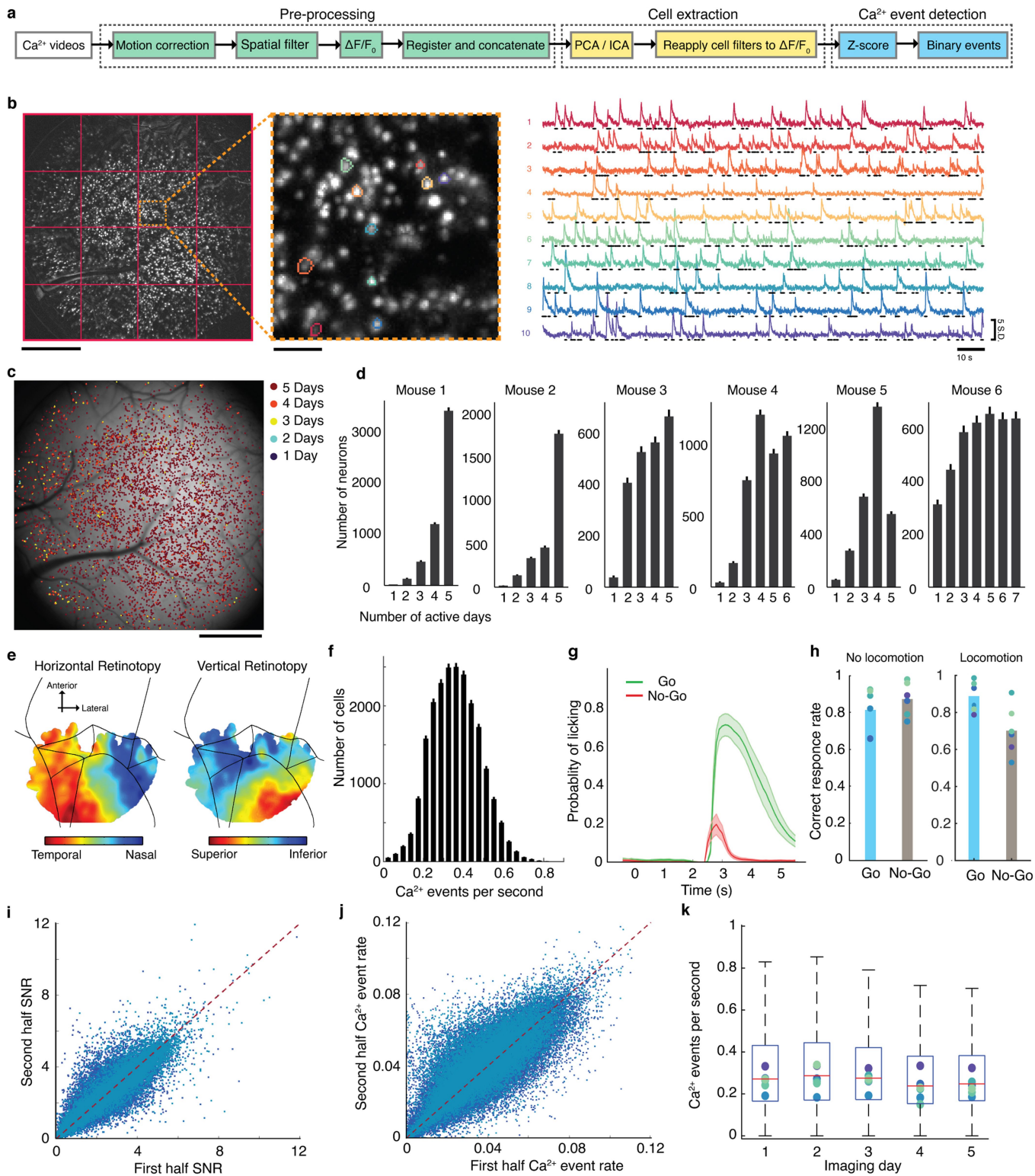
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04724-y>.

Correspondence and requests for materials should be addressed to Sadegh Ebrahimi or Mark J. Schnitzer.

Peer review information Nature thanks Tatiana Engel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

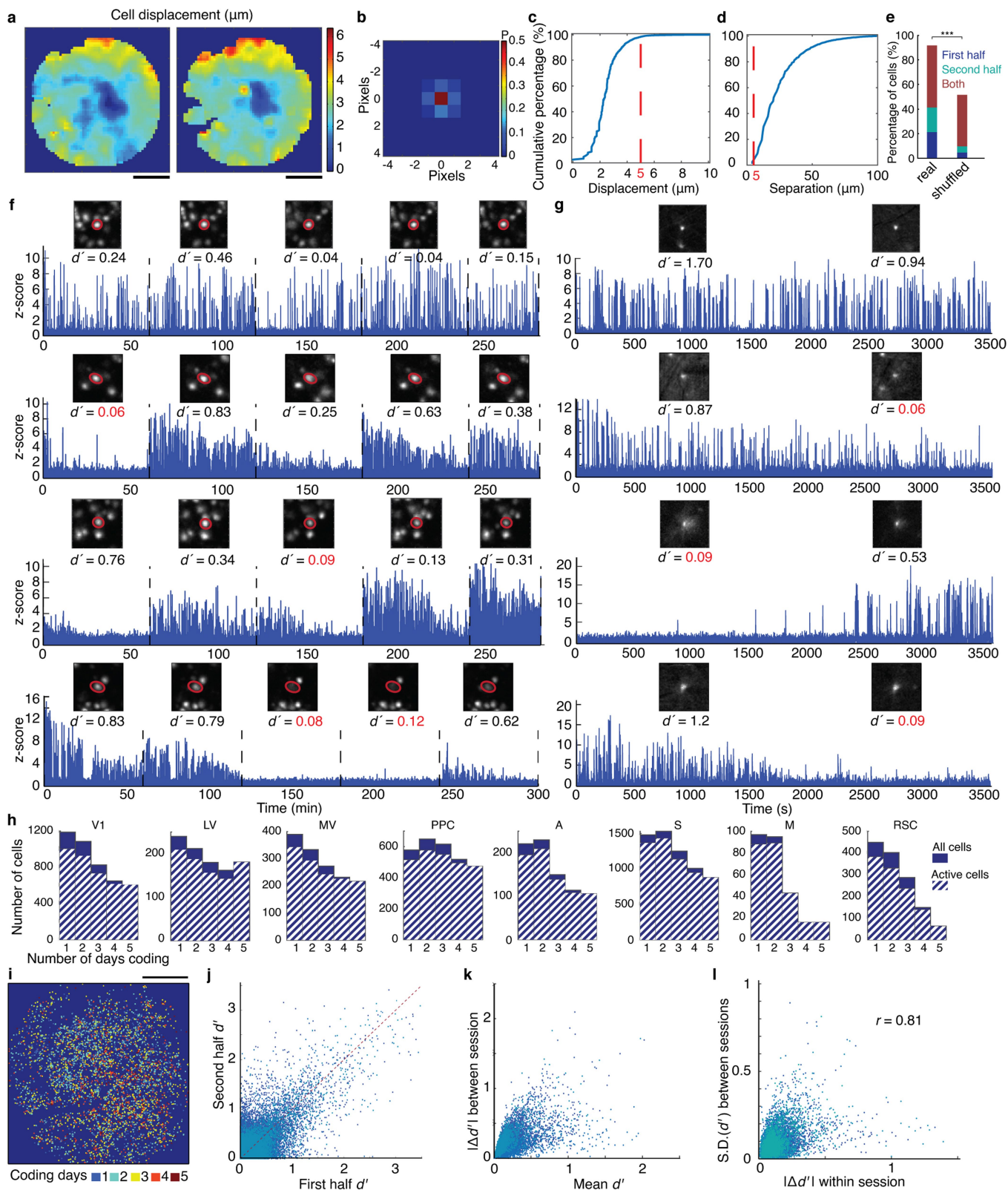


Extended Data Fig. 1 | See next page for caption.

Article

Extended Data Fig. 1 | Long-term imaging and computational analysis of neural Ca²⁺ dynamics across multiple cortical areas during a visual discrimination task. (a) Schematic of the algorithmic pipeline used for video preprocessing and cell extraction, as implemented using cluster computing. **Pre-processing** (steps shown in green): For each movie of Ca²⁺ activity, we performed an image registration across all frames of the movie to correct for small displacements of the brain. We removed background noise and neuropil Ca²⁺ activity by applying a spatial Gaussian high-pass filter ($\sigma = 80 \mu\text{m}$) and computed a movie of the relative changes in fluorescence ($\Delta F(t)/F_0$). We then aligned and concatenated all the $\Delta F(t)/F_0$ movies for each individual mouse, across all imaging sessions. **Cell extraction** (steps shown in yellow): We divided each concatenated movie into 16 spatial tiles and then extracted individual cells within each tile by successively applying principal component and independent component analyses (PCA/ICA algorithm)⁵⁶ to all tiles in parallel using the Stanford Sherlock computing cluster (using up to 320 cores and -2 TB of memory for each concatenated movie). **Ca²⁺ event detection** (steps shown in cyan): We converted the $\Delta F(t)/F_0$ traces for each neuron to traces expressing the time-dependent fluorescence changes as a z-score, $z(t)$, relative to the s.d. of the baseline fluctuations in each cell's fluorescence trace (computed separately for each imaging session). We detected Ca²⁺ events by identifying Ca²⁺ transients that attained a peak fluorescence value of $z(t) \geq 4$ s.d., and we assigned the cell as being 'active' within the interval between the initial threshold crossing and the time at which the Ca²⁺ event attained its peak fluorescence (Methods). (b) *Left*: A maximum projection image over an entire concatenated set of Ca²⁺ movies from an example mouse. Red lines mark the 4×4 set of tiles that we processed in parallel during cell extraction. Scale bar: 1 mm. *Middle*: Magnified view of the area enclosed in orange in the left panel. Scale bar: 0.1 mm. *Right*: Z-scored traces (coloured traces) of fluorescence Ca²⁺ activity for 10 example neurons in the middle panel marked with colour-corresponding boundaries. Raster traces show the binarized patterns of activity for each cell. (c) Most detected cells were active in all recording sessions, as illustrated via a map, computed for one example mouse, in which each detected cell is marked with a colour-code indicating the number of days

in which it was detected as active (Methods). Scale bar: 1 mm. (d) Histograms of the number of days that each cell was detected as active for 6 different mice. Error bars are s.d. estimated as counting errors. (e) Vertical and horizontal retinotopic maps of visual cortex in an example mouse (Methods). After identifying borders of area V1 determined by retinotopic mapping studies in each mouse, we aligned these borders with those in the Allen Brain Observatory map of the mouse cortex and thereby inferred the locations of other brain areas. (f) Histogram of the mean Ca²⁺ event rate for each of 21,570 cortical neurons ($N = 6$ mice). Error bars are s.d. estimated as counting errors. (g) Mean probability of licking within individual time bins (0.1 s duration) over the time course of a trial, averaged over all trials and trained mice, for Go (green) and No-Go (red) trials. Shaded areas denote s.e.m. over $N = 6$ mice. After mice learned to discriminate between Go and No-Go visual stimuli, we trained them to withhold licking behaviour during the stimulus presentation, [0 s, 2 s], and delay, [2 s, 2.5 s], intervals and to respond only during the response interval, [2.5 s, 5.5 s] (Fig. 1 and Methods). Trained mice occasionally licked before the response interval; we discarded these trials from our analyses to allow inferences regarding stimulus encoding, decision-making, and motor preparation in the absence of overt licking responses. (h) The mean behavioral performance of all mice on Go (cyan) and No-Go (grey) trials in which the mouse did (*right*) or did not exhibit locomotor behaviour (*left*) (Methods). Individual data points denote values from individual mice. (i, j) For every individual cell (blue data points), the plots show the mean signal-to-noise ratio (SNR) of Ca²⁺ activity, **i**, or the mean occurrence rate of Ca²⁺ transient events per time bin (0.1 s duration), **j**, in the first half of each imaging session versus that in the second half of the same session. From linear regression, the mean SNR and Ca²⁺ event rate in the second half of each session were $96 \pm 2\%$ ($N = 6$ mice) and $99 \pm 3\%$ ($N = 6$ mice), respectively, of their values in the first half. (k) A box and whisker plot of the Ca²⁺ event rate across all cells imaged for 5 days in each mouse ($N = 2236$ -5292 cells). Horizontal lines indicate median values across all cells, boxes cover the second and third quartiles, and whiskers extend to 1.5 times the interquartile distance. Dots show median values for individual mice.

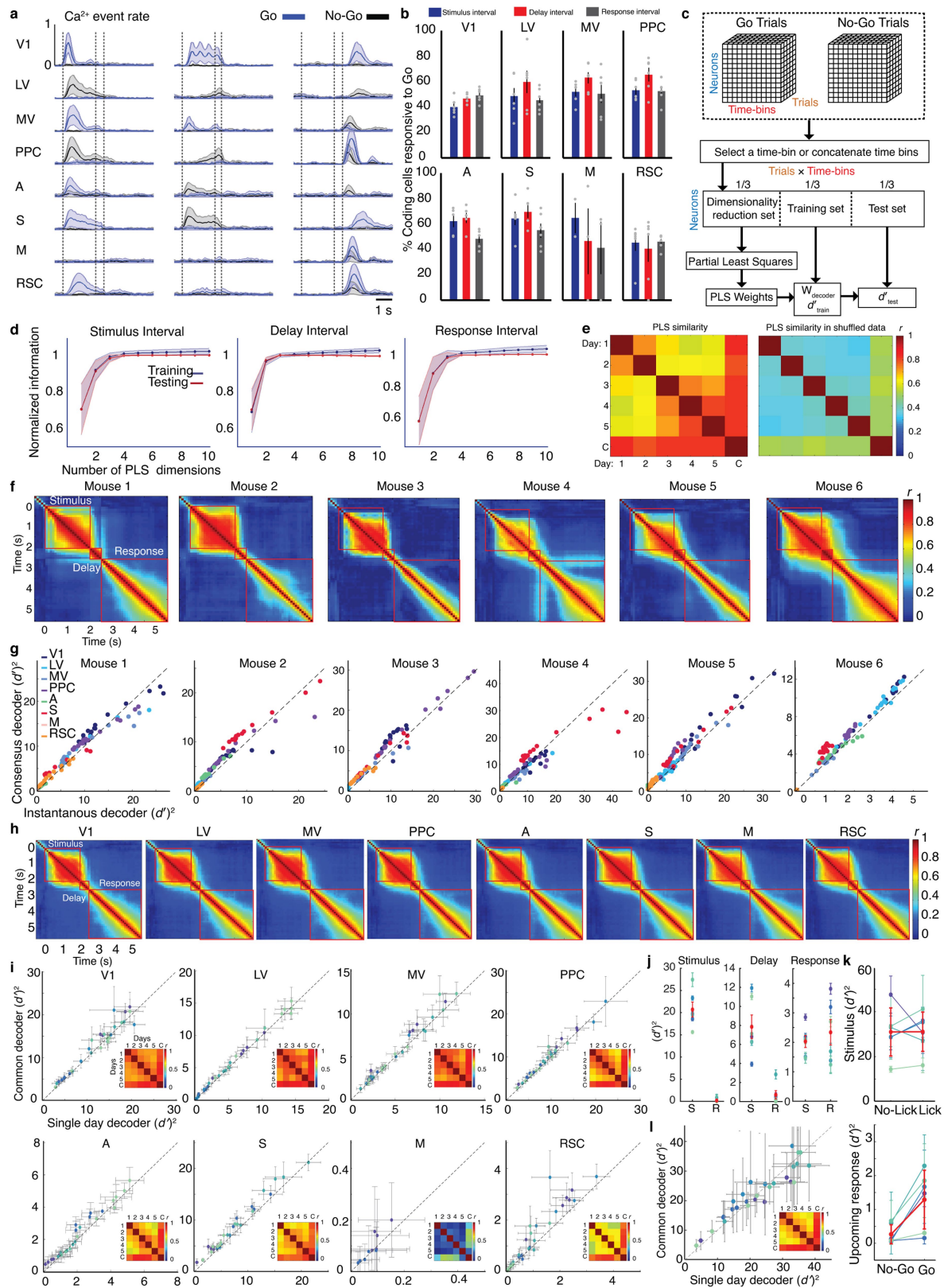


Extended Data Fig. 2 | See next page for caption.

Article

Extended Data Fig. 2 | Cortical neurons exhibit variable coding properties across timescales from minutes to days. (a) Maps for two example mice, showing how mean lateral displacements in cells' centroid positions across multiple imaging sessions depended on their locations in the field of view. Across most of the field of view, mean displacements were <1 pixel, corresponding to <4 μm . To determine these displacements, we first computed the maximum projection image (MPI) of the Ca^{2+} video from each session. Using the MPI from the first session as a reference, we computationally aligned it to the MPI from each of the other sessions. We then computed the spatial cross-correlation function between patches of the MPI containing ≥ 10 cells from the first session (patch size: $256 \mu\text{m} \times 256 \mu\text{m}$) and MPIs from each of the other sessions. For each session other than the first, we determined the displacement of a patch to be the argument of the cross-correlation function that yielded its maximum value. We averaged these displacements across all sessions after the first. By examining all possible MPI patches (spaced $64 \mu\text{m}$ apart) in this way, we created the maps shown. Scale bars: 1 mm. (b) Two-dimensional probability distribution of cells' daily lateral displacements from their mean position, averaged across all days and imaged neurons (21,570 cells) from $N = 6$ mice (Methods). About 50% of the time, cells had a displacement of zero pixels from their mean position; 98.5% of the time these displacements were ≤ 1 pixel ($4 \mu\text{m}$). (c) Cumulative distribution of cells' mean displacements (averaged over all sessions) from their mean positions across the entire experiment. Red dashed line indicates that 95.4% of cells had a mean displacement of $\leq 5 \mu\text{m}$. (d) Cumulative distribution of the lateral separations between nearest neighbour cell pairs. Red dashed line indicates that only 2% of nearest neighbour cell pairs were within $5 \mu\text{m}$ of each other. (e) Among 18,528 cells with significant d' values on one or more sessions for encoding the trial-type in the stimulus period ($P < 0.01$; permutation test; $N = 94$ –354 trials), 41% of these had significant d' values in only one half-session, split nearly evenly between the first (21%) and second (20%) half-sessions. Whereas in trial-shuffled data, only 10% of the cells had this variable coding, a highly significant difference from the real data ($P < 0.001$) indicating that trial-shuffling diluted the temporal concentration of trials in which cells had coding responses. In real data, 91% of the 18,528 cells retained significant coding in one or both halves of the full sessions in which they displayed significant coding ($P < 0.01$; permutation test; 40–175 trials). But in trial-shuffled data, only 51% of the cells retained this coding in one or both half-sessions, a highly significant difference from real data ($***P < 0.001$; permutation test; 94–354 trials), again showing that in real data cells had temporally concentrated coding epochs far more than expected by chance. All s.d. values on the above percentages of cells were estimated as counting errors

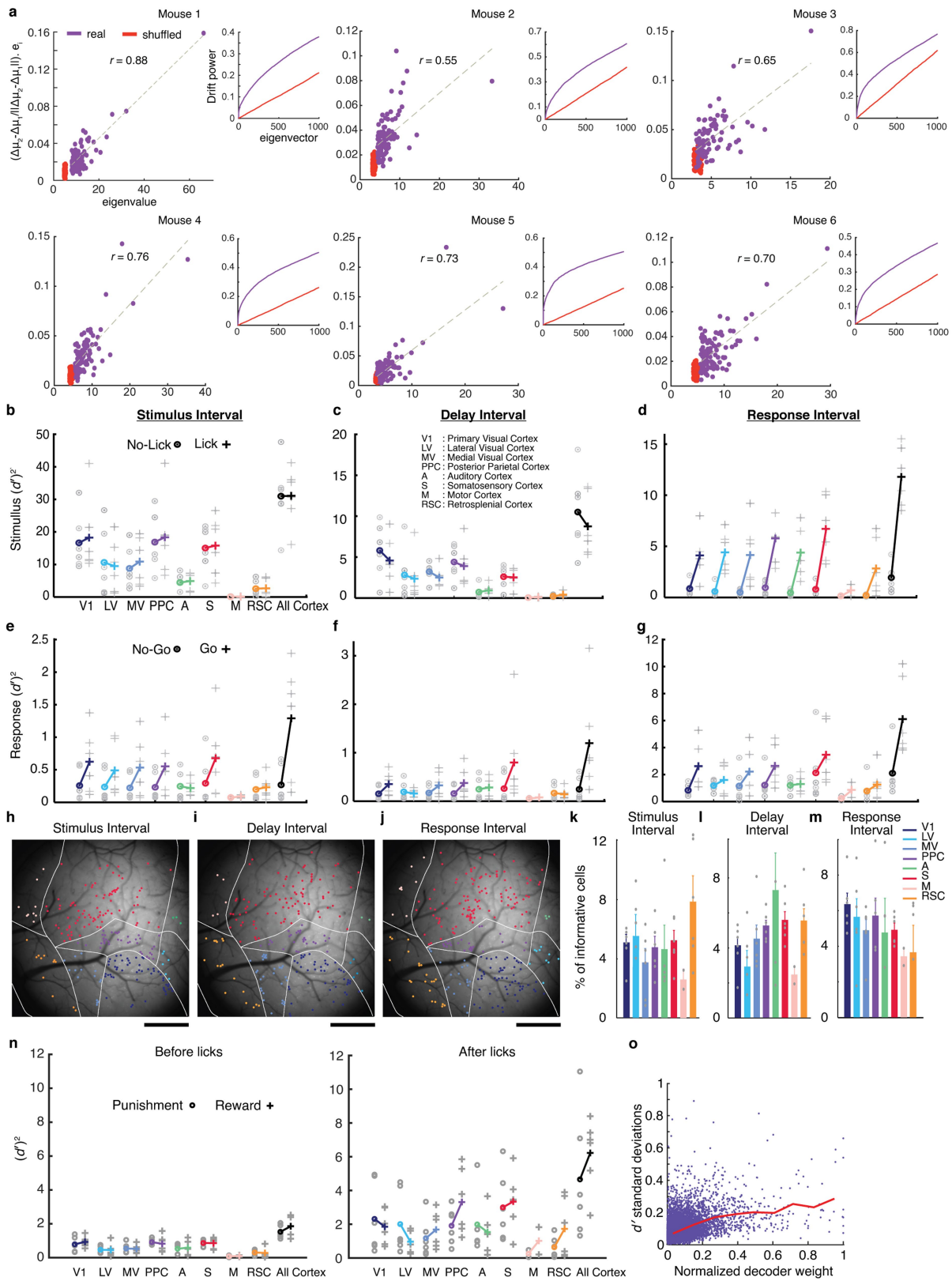
and were 0.1–0.4%. (f) Some neurons had coding properties that varied across days. For 4 example cells (from areas PPC, MV, V1 and PPC, from top to bottom), shown are traces of the neuron's fluorescence intensity (z-scored values of $\Delta F/F_0$) as a function of time across 5 imaging sessions. Vertical dashed lines mark transitions between successive sessions. Insets show maximum projection images of the example neurons from each session. Values of d' denote the fidelity with which one can distinguish the two visual stimuli based on the binarized event train of the cell's Ca^{2+} activity (Methods). In f and g, values of d' coloured red are those for which the stimuli cannot be significantly distinguished, as determined using a permutation test over the set of stimulus trials and requiring $P < 0.01$ for significance. (g) Some cells had coding properties that varied within the 1-h sessions. Shown are fluorescence traces (z-scored values of $\Delta F/F_0$) for 4 cells (from areas LV, V1, MV and LV, from top to bottom) as a function of time across one session. We measured d' values of single neurons for distinguishing the two visual stimuli during the first and second halves of each session based on their binarized Ca^{2+} event traces. Neurons that actively fired across the session exhibited variability in their visual coding, as did cells that were active during only a portion of the imaging session. Insets: Example Ca^{2+} event images show that the same cells were imaged in the first and second halves of each session. (h) Histograms showing the numbers of days that neurons from each area significantly encoded the visual stimulus-type (permutation test over the set of stimulus trials; requiring $P < 0.01$ for significance), for all cells that did so in at least one session (solid bars) and for the subset of these cells with statistically significant levels of Ca^{2+} activity in every imaging session (hashed bars). (i) Map of neurons from an example mouse, with each cell's colour denoting the number of days the cell significantly encoded the visual-stimulus type. Cells with different day-to-day reliabilities of stimulus-encoding were interspersed across the field of view. Scale bar: 1 mm. (j) Scatter plot in which, for every individual cell (blue data points), the d' value for stimulus discrimination during the first half of each imaging session is plotted against the d' value for the second half of the same session. (k) Scatter plot in which, for every cell (blue data points), the mean d' value for stimulus discrimination (averaged over all sessions) is plotted against the range of d' values determined for the same cell across all sessions. (l) Scatter plot in which, for every cell (blue data points), the mean difference between the d' values for stimulus discrimination determined for the first and second halves of each session is plotted against the s.d. of the d' values determined for the same cell across all sessions. Variability in d' values within a session was highly correlated ($r = 0.81$) with variability across sessions, suggesting some neurons have greater intrinsic variability in the fidelity of stimulus encoding than others.



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Neural ensemble representations of the visual stimuli were invariant over most of the stimulation period. (a) Mean time-dependent occurrence rates of Ca^{2+} transient events per time bin (0.1s duration) across different intervals of the trial-structure (demarcated by vertical lines) for 24 neurons from 8 cortical areas, averaged across 5 sessions in one mouse on Go (blue traces) and No-Go (black) trials. Shading: s.d. across 415 trials of each type. (b) For cells that responded significantly to one of the two stimuli (see Fig. 2c), shown are mean percentages of coding cells responding to the Go stimulus in each of 8 areas. The remainder of coding cells responded to the No-Go stimulus. Error bars: s.d. across $N = 6$ mice. (c) Procedure for training cross-validated instantaneous or consensus linear decoders. After constructing an unbiased dataset with equal numbers of Go and No-Go trials, we divided the set of trials into 3 equal portions, one for dimensionality reduction, another for decoder training, and the third for decoder testing. Using the first subset, a partial least squares (PLS) analysis identified a low-dimensional subspace of the neural ensemble activity that was informative for stimulus discrimination. Within this subspace, we used the second subset of trials to train a Fisher linear decoder (indicated by the vector $\mathbf{W}_{\text{decoder}}$). We used the third subset to test decoder performance. For training and testing datasets, we computed the fidelity, d' , with which the stimuli could be distinguished. To train decoders for identifying the mouse's decision, we followed the same procedures, starting with equal numbers of correctly and incorrectly performed trials with a given stimulus. (d) We trained consensus decoders during the stimulus, delay, and response intervals of correctly performed trials. Plots show mean $(d')^2$ values for decoder training (blue) and testing (red) datasets, versus the number of PLS dimensions used. When constructing each decoder, we used the number of PLS dimensions that maximized $(d')^2$ for testing datasets. All plotted $(d')^2$ values are separately normalized for each mouse to the maximum $(d')^2$ determined with the testing data. On average, with >5 PLS dimensions decoders overfit the training data, yielding greater $(d')^2$ values than for testing data. For shuffled datasets, 1 or 2 PLS dimensions yielded maximal $(d')^2$ values (data not shown). Shading: s.d. across $N = 6$ mice. (e) We determined the similarity of the subspaces defined by the top 3 PLS dimensions for each mouse on different days (1–5) or for its across-day, common decoder (C) (Methods). We used the top 3 dimensions, since these contain most of the information (d). The matrices show mean similarity values for all pairs of subspaces, averaged over $N = 6$ mice, for real (left) and shuffled (right) datasets. For real datasets, PLS dimensions for common decoders were highly similar to those for single-day decoders. (f) Optimal linear decoders of stimulus-type retained a constant form across stimulus presentation. Plots show Pearson correlation coefficients, r , between all pairs of instantaneous decoders (constructed using all imaged neurons in each mouse), for each time bin. (g) Due to the stationarity of the optimal decoders across stimulus

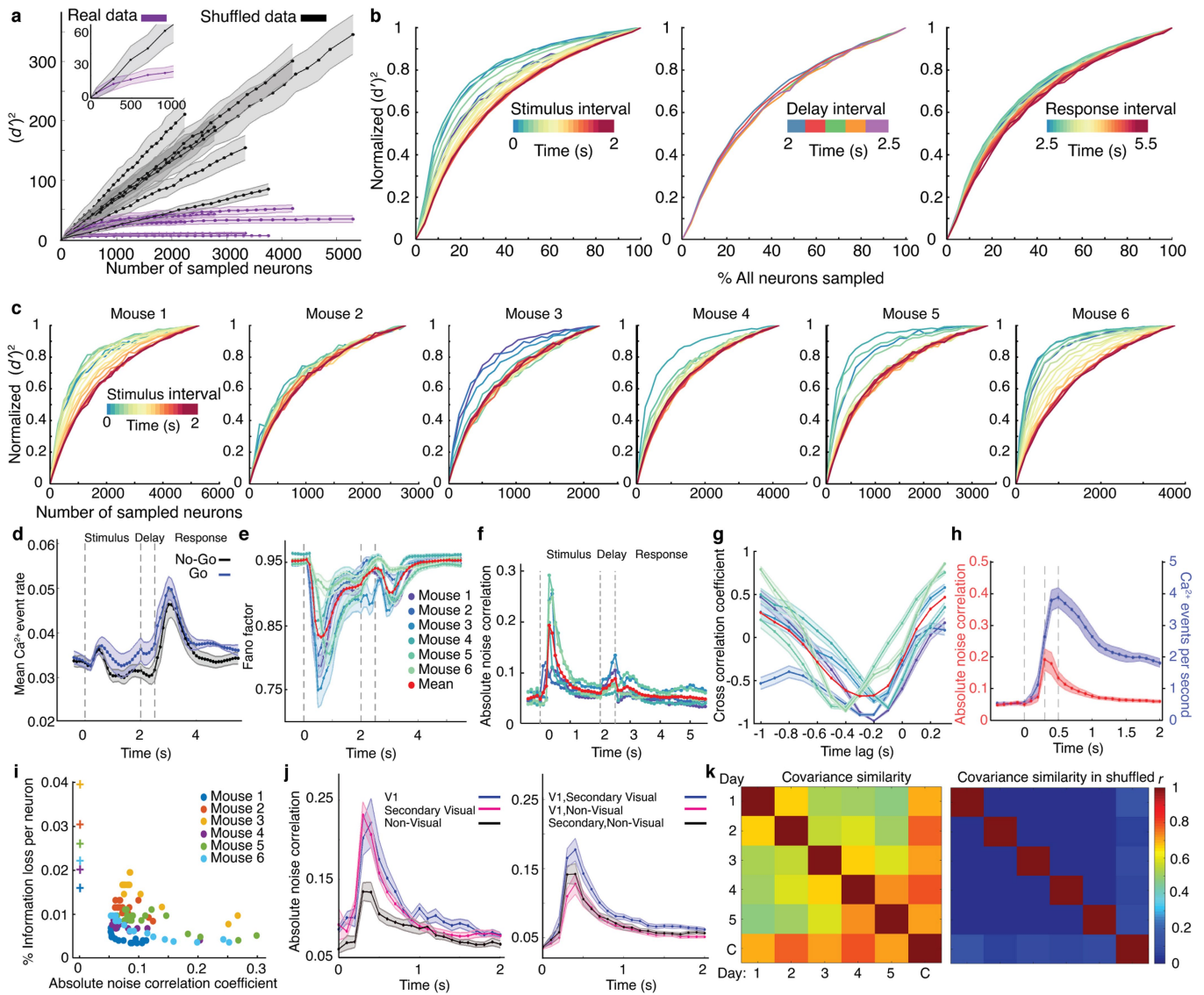
presentation, f, consensus and instantaneous decoders performed nearly equivalently. Plots show mean $(d')^2$ values for consensus decoders of stimulus-type versus those for instantaneous decoders, for correctly performed trials. Each datum shows testing results attained by applying each decoder-type to data from one time bin during stimulus presentation. In some mice, e.g. Mice 5 and 6, the consensus decoder achieved slightly superior performance, presumably due to its larger training dataset. (h) Similar results to those of f, computed for different areas and averaged over 6 mice. (i) Similar results to those of Fig. 3c, computed separately for different areas. (j) To measure the information captured by trial-type decoders about the stimulus (S) or mouse's response (R) in the stimulus (left), delay (middle) or response (right) periods, we projected neural activity on all 4 trial-types (Hit, Miss, Correct Rejection, and False Alarm) onto the common trial-type decoders trained for each period using correctly performed trials. We then computed $(d')^2$ values using trials in which either the stimulus or response was held constant but the other varied. Information, $(d')^2$, about the stimulus did not vary significantly between Lick and No-Lick trials, so we averaged $(d')^2$ values across both stimuli in each plot's left column. Response-coding was much stronger on Go trials (see k), so right columns only show $(d')^2$ values from Go trials. Each blue point shows data from one mouse (mean \pm s.d., $N = 100$ different subsets of trials, each with equal numbers of trials of the two types). Red points: averages across mice (mean \pm s.e.m., $N = 6$). During the stimulus period, common decoders nearly exclusively captured stimulus information, which was 691 ± 315 times greater (mean \pm s.e.m.; $N = 6$ mice) than information captured about the mouse's response. In the delay period, response information rose. During the response period, common decoders captured response information that was comparable or greater to information about the stimulus. (k) The mean Fisher information encoded by neural ensemble activity about the stimulus-type was independent of the mouse's response (top), as $(d')^2$ values for consensus common stimulus decoders trained and tested on 'No-Lick' trials were indistinguishable to those for 'Lick' trials ($P < 0.7$; Wilcoxon signed-rank test; $N = 6$ mice). On 'Go' but not 'No-Go' trials, the upcoming response could be predicted ($P < 0.01$; permutation test; $N = 40 - 754$ trials) from neural activity during stimulus presentation (bottom). For each comparison, we constructed training datasets for the two decoders to have equal numbers of trials, 50% of each type. Blue-shaded points are from individual mice; error bars are s.d. ($N = 100$ different randomly chosen sets of trials. Red points are means; error bars are s.d. ($N = 6$ mice). (l) Control analysis for Fig. 3c. Across-day common consensus decoders performed equivalently to single-day consensus decoders when they were trained with equally sized datasets. Here we trained common decoders by sub-sampling trials from each session so the training dataset had the same of number of trials as that of the day with the smallest number of trials. We trained single-day decoders using this same number of trials.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Neural ensemble representations of the visual stimuli and the mouse's response were widespread across neocortical areas. (a) Plots analogous to Fig. 3f for individual mice. In all 6 mice, day-to-day changes in coding were correlated with within-day, trial-to-trial fluctuations ($r = 0.85, 0.66, 0.79, 0.76, 0.83, 0.76$ and P was between 5×10^{-14} – 5×10^{-29} for mice 1–6 for the real datasets, but $0.1 \leq r \leq 0.15$ and $0.12 \leq P \leq 0.92$ for trial-shuffled datasets). (b) We trained consensus common decoders to discriminate the visual stimuli based on neural activity either in individual cortical areas or across the visible cortical regions, during stimulus presentation on 'No-Lick' trials (trials when the mouse withheld licking responses) and Lick trials (on which the mouse made a licking response). Thus, decoders for 'No-Lick' trials discriminated 'Correct Rejection' from 'Miss' trials, and decoders for 'Lick' trials discriminated 'Hit' from 'False Alarm' trials. Both decoder types were trained on equally sized datasets, with equal numbers of trials of each type. We evaluated decoder performance across individual time bins and then plotted the maximum (d')² values attained for each mouse across all time bins during stimulus presentation (0.5–2 s after stimulus onset). (d')² values for stimulus decoding were independent of the mouse's upcoming 'Lick' or 'No-Lick' response ($P < 0.7$; Wilcoxon signed-rank test, $N = 6$ mice). Across b–g, grey and coloured symbols respectively denote (d')² values for individual mice and mean values averaged over $N = 6$ mice; note that y-axis scales vary substantially across the graphs. (c, d) Using the methods of b, we trained consensus common decoders to discriminate the visual stimuli using neural activity in different cortical areas during the delay (c) and response (d) periods. Similarly to b, we evaluated decoder performance across individual time bins and plotted the maximum (d')² values attained for each mouse across all time bins during the delay, c, or response, d, periods. Whereas (d')² values for stimulus decoding during the delay period were independent of the mouse's upcoming response ($P < 0.3$; Wilcoxon signed-rank test; $N = 6$ mice), response period (d')² values were significantly greater for 'Lick' trials ($P < 0.03$). The latter, higher values of (d')² could stem from the divergent neural signals evoked by receipt of a reward or air puff on 'Hit' and 'False Alarm' trials, respectively. (e–g) Using methods analogous to those in b, we trained consensus decoders of the mouse's response on 'Go' and 'No-Go' trials based on neural activity in different cortical areas during the stimulus (e), delay (f), and response (g) intervals. As in b–d, we evaluated decoder performance across individual time bins and constructed plots using the maximum (d')² values attained for each mouse across all time bins during the stimulus (0.5–2 s after stimulus onset), e, delay, f, or response g, periods. To determine neural representations of the mouse's response during the response interval, g, we used data across the full 3-s response interval. Within this interval, the mouse received rewards and

aversive air-puffs at variable time points. Thus, a distinct analysis would be needed to separate coding relating to receipt of the rewarding and aversive stimuli from that relating to the mouse's actions. (d')² values for response decoding were greater for 'Go' trials during stimulus presentation ($P < 0.03$; Wilcoxon signed-rank test; $N = 6$ mice), delay ($P < 0.06$), and response ($P < 0.06$) intervals. These higher (d')² values could reflect signals associated with reward prediction, motor planning and action on correctly performed 'Go' trials. (h–j) Map of the cortex for the mouse of Fig. 3g–j. Coloured dots mark locations of cells with the greatest contributions to the response decoder score (defined as cells with decoder weights deviating > 2 s.d. from mean values) during stimulus (h), delay (i), and response (j) intervals. Because the mouse's response was only weakly encoded in the neural dynamics on 'No-Go' trials (see e–g), we created h–m based on the response decoders found by analysis of 'Go' trials. Cells in each brain area are coloured differently, following the colour scheme in m. Scale bars: 1 mm. (k–m) Mean \pm s.e.m. ($N = 6$ mice) fractions of neurons in each area that had response decoder weights deviating > 2 s.d. from mean values, during the stimulus (k), delay (l), and response (m) intervals. (n) Right, We measured the information (d')² conveyed about reward and punishment in each area by studying the neural activity evoked when the mouse licked. To evaluate the encoding of punishment, we compared the mean neural ensemble activity in the first 0.5 s after licks that were punished with air-puffs versus after licks that occurred during timeout periods and that elicited neither punishment nor reward. To evaluate reward encoding, we compared the mean neural ensemble activity in the first 0.5 s after licks that occurred during timeouts versus after licks triggering a reward. Both punishment and reward were represented to varying extents across the different areas. Notably, these representations could relate to any aspect of the rewarding or aversive experience, such as the experience of receiving or blinking in response to an aversive air-puff or receiving or tasting a reward. Left, As a control, we performed the same calculations as for the right panel but using neural activity that occurred within the 0.5 s intervals just before licks. As expected, during these periods there was notably less information encoded about upcoming rewards or punishments than about rewards or punishment the mouse has just received. (o) A graph of the s.d. of (d')² values for each cell (individual data points) across all days of the study, for every cell with a significant ($P < 0.01$) d' value for trial-type encoding on at least one day, as a function of the cell's weight in the across-day common decoder. Decoder weights are normalized by the maximum weight in each mouse. Results show that cells can have stable or variable coding properties, irrespective of their decoder weights. Nevertheless, coding variability generally increases for cells with larger weights, as shown by the red line, which plots the mean s.d. in (d')² values, averaged over all cells within x-axis bins of 0.1.



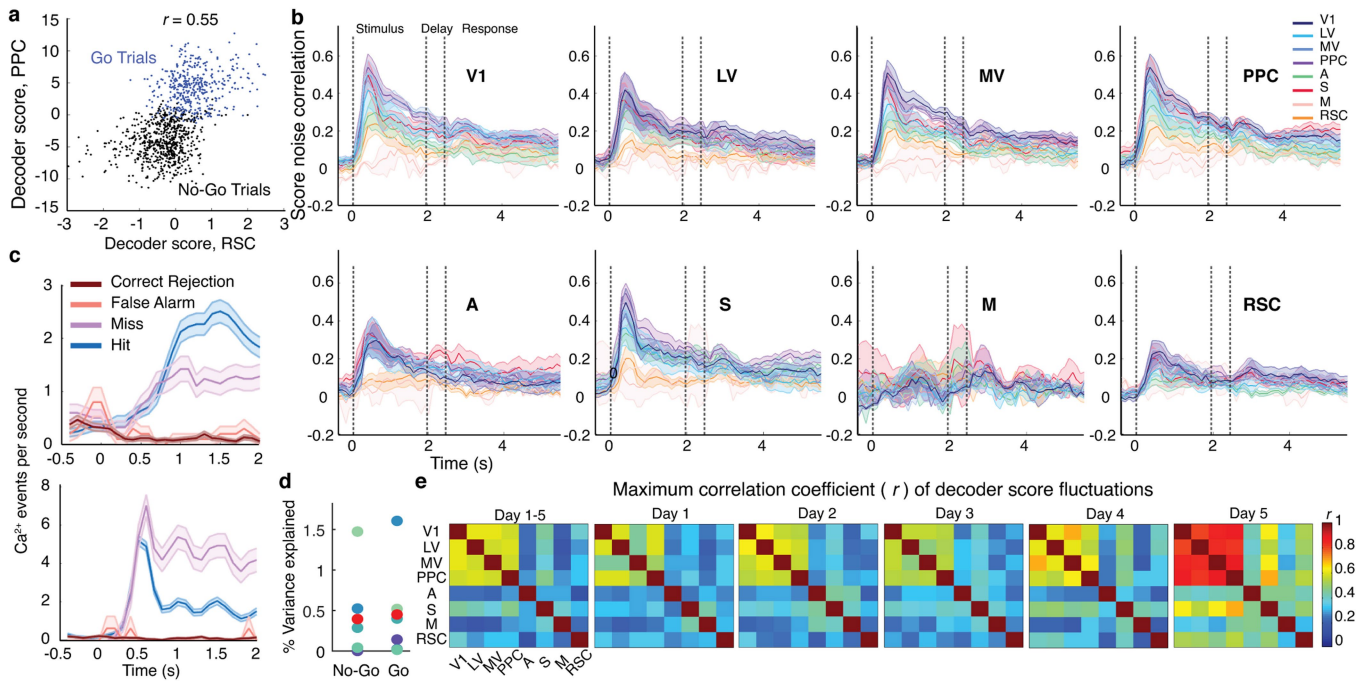
Extended Data Fig. 5 | See next page for caption.

Article

Extended Data Fig. 5 | Information-limiting noise correlations and coding redundancy peaked just after stimulus onset and then declined for the rest of stimulus presentation.

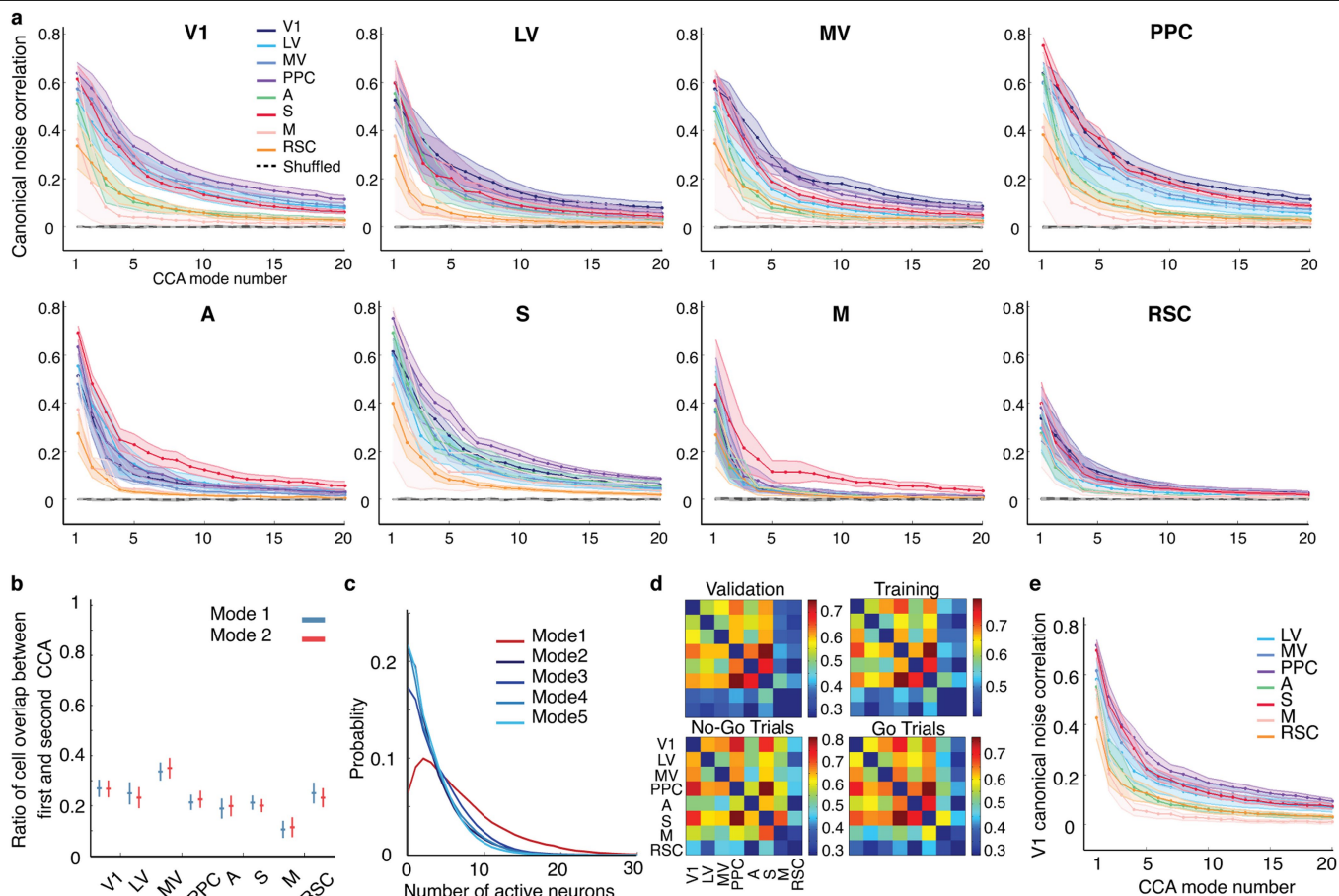
(a) The fidelity with which stimulus identity could be decoded from neural ensemble activity saturated for large (>2000) populations of cells, for real (purple curves) but not trial-shuffled (black curves) datasets. To study ensembles of each size (x -axis), we randomly chose 100 different subsets of cells from the entire pool of neurons imaged across all areas. We then trained and tested optimal linear Fisher decoders using the neural activity within [0.4 s, 0.5 s] after stimulus onset on correctly performed trials. We quantified decoding performance using $(d')^2$, which relates to the Fisher information the neural dynamics conveyed about the trial-type. Each curve shows data from one mouse. Whereas $(d')^2$ saturated for large neural populations in real data, this did not occur for trial-shuffled datasets in which correlated fluctuations were scrambled. Shading: s.d. across all 100 subsets of cells for each ensemble size. *Inset:* Magnified view near the graph origin for one mouse. **(b)** Using the methods of **a**, we assessed how well optimal linear decoders could discriminate Go and No-Go trials. Plots show mean $(d')^2$ values for this discrimination as a function of neural ensemble size and for different time bins, averaged over $N = 6$ mice. The size of the cell ensemble at which $(d')^2$ saturated rose substantially with time during stimulus presentation but stayed relatively constant during the delay and response periods. $(d')^2$ values are normalized relative to their maximum (saturating) value at each time bin. Ensemble size values are normalized relative to the total number of cells recorded in each mouse. **(c)** Plots like those of **b**, for individual mice during stimulus presentation. Data are shown only for time bins in which $(d')^2$ values were significantly greater than for control datasets in which the trial-type labels were randomly shuffled ($P < 0.01$; permutation test; $N = 710$ –1340 trials). **(d)** Mean \pm s.e.m. ($N = 6$ mice) Ca^{2+} event rates per time bin (0.1 s duration) for all neurons on correctly performed Go and No-Go trials. These event rates had near identical time dependencies on trials of the two types, but the temporal variations were distinct from those of decoder score fluctuations (Fig. 4b) or correlated fluctuations in cells' dynamics (panel **f**). Dashed vertical lines in **d**–**f** demarcate stimulus, delay and response periods of the trial structure. **(e)** Time dependence of the mean Fano factor, determined for each mouse by computing for each cell the ratio of the variance in the cell's Ca^{2+} event rate to its mean Ca^{2+} event rate, on correctly performed trials. Shading: s.e.m. values ($N = 2236$ –5292 cells). Legend also applies to **f** and **g**. **(f)** Noise correlations between pairs of cells with similar stimulus tuning rose sharply after stimulus onset, peaked ~ 0.2 s after stimulus onset, and then decayed to baseline values. Each coloured trace shows the mean absolute value of noise correlation coefficients for all pairs of similarly tuned cells across all imaged areas in each mouse. **(g)** Cross-correlation functions between the dynamics of absolute

noise correlations across pairs of cells, shown in **f**, and the Fano factor, shown in **e**, as determined for each mouse over the 2-s-stimulus period. The graph shows that changes in pairwise noise correlation coefficients were negatively correlated with and most predictive of upcoming variations in the Fano factor with a lead time of ~ 200 ms. Shading: s.e.m. values ($N = 10$ –20 time bins for each abscissa value). **(h)** Plot of the mean time-dependent rate (blue trace) of Ca^{2+} events in Go-stimulus-tuned neurons on Go trials and No-Go-stimulus-tuned neurons on No-Go trials, averaged over both cell-types and $N = 6$ mice. Also shown is the mean absolute noise correlation coefficient (red trace) for pairs of similarly tuned neurons, computed as in **f** for the same 6 mice. Notably, changes in noise correlation coefficient levels peaked sooner after stimulus onset than Ca^{2+} activity rates of tuned cells. After reaching their peak values, noise correlation coefficients declined back to baseline values by stimulus offset, whereas Ca^{2+} activity rates did not. These differences make it hard to explain the dynamics of noise correlation coefficients as resulting simply from changes in neural activity rates. Shading: s.e.m. across 6 mice. **(i)** Plot showing the change in information encoded by the neural ensemble if one cell were to become silent, assessed using instantaneous decoders. Each dot denotes the result from one time bin. (As shown in **c** and **f**, noise correlation coefficients vary with time following stimulus onset). Results for trial-shuffled data, in which correlated fluctuations were scrambled, are denoted with crosses and reveal a greater sensitivity to loss of one neuron. **(j)** *Left*, Traces of mean absolute noise correlation coefficients as a function of time during stimulus presentation, determined as in **f** for cell pairs in primary visual cortex (V1; blue trace), secondary cortical visual areas (areas LV, MV and PPC; red trace) or non-visual cortical areas (areas A, S, M and RSC; black trace). *Right*, Traces of mean absolute noise correlation coefficients between pairs of coding neurons located in different brain areas. The rise in noise correlations for similarly tuned cells in visual cortex is greater than that for cells outside visual cortex ($P < 0.03$; Wilcoxon signed-rank test; $N = 6$ mice). Shading: s.e.m. across $N = 6$ mice. **(k)** We calculated the covariance in the neurons' responses on each trial-type and on each day. We then averaged the covariance matrices for the two trial-types and computed the top 3 eigenvectors for each day. *Left*, A plot showing the similarity between the pairs of different subspaces (Methods), each defined by the top 3 eigenvectors of the noise covariance matrix on each day of experimentation. The matrix row and columns labelled 'C' is for the noise covariance matrix computed for the set of all trials across all days. *Right*, As control, we computed the subspace similarities for trial-shuffled datasets in which each neuron's responses were permuted across trials with the same stimulus. Overall, noise covariance structure in the real data was significantly similar across days, to a degree much beyond that in shuffled datasets.



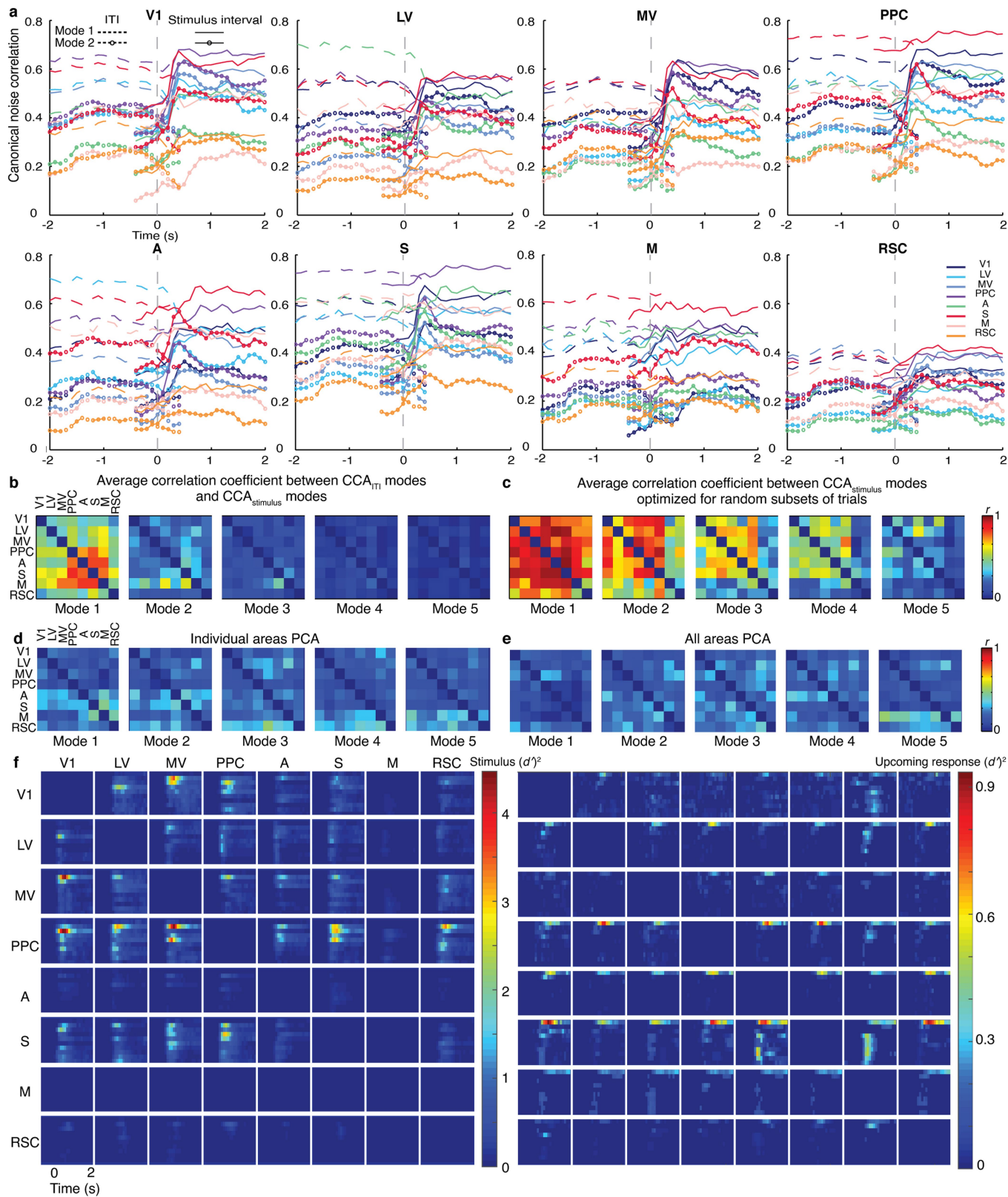
Extended Data Fig. 6 | The discriminability of the two stimuli based on their evoked neural dynamics fluctuated trial-by-trial in a way that was highly correlated between cortical areas. (a) Example scatter plot for an individual mouse in which the instantaneous stimulus decoder scores based on the activity patterns of cortical area PPC are plotted against those for cortical area RSC. Each data point shows results for an individual trial, at 0.5 s after stimulus onset, for Go trials (blue data points) or No-Go trials (black data points). Stimulus decoder scores for the two brain areas exhibit positively correlated trial-to-trial fluctuations. (b) Traces showing the mean time-dependent correlations of the fluctuations in instantaneous stimulus decoder scores for 8 different cortical areas and each of the other 7 brain areas within the imaging field-of-view. For most pairs of brain areas, these correlated noise fluctuations in decoder scores attained their maximum shortly after stimulus onset and then gradually decayed. Decoder training and testing was limited in this analysis to trials that the mice performed correctly. Shading: s.e.m. over $N = 6$ mice. Vertical dashed lines demarcate the stimulus presentation, delay and response intervals. (c) Two plots showing examples of stimulus-coding cells whose responses were modulated by the mouse's response. Each plot shows the mean rate of Ca^{2+} events in an individual neuron, as a function of time relative to stimulus onset at $t = 0$, for the 4 different trial-types. The cell of the top plot is from area MV, and the cell of the bottom plot is from PPC. Both cells had P -values of < 0.01 for stimulus-coding on Lick and No-Lick trials, and also

had $P < 0.01$ for response-coding on GO-trials). We determined P -values through comparisons to trial-shuffled datasets (1000 different sub-samplings and random permutations of trials using equal numbers of trials of both stimulus- or response-types). The separation between the traces for Hit and Miss trials shows the extent of response-related modulation on trials with a Go stimulus. Shading: s.e.m. over trials (410 Hit trials, 218 Miss trials, 665 Correct Rejection trials, 100 False Alarm trials). (d) To determine if the elevated correlated noise fluctuations along the stimulus-coding direction within the interval $[0.2 \text{ s}, 0.5 \text{ s}]$ after stimulus onset (when correlations were at their peak) reflects choice information relating to the formation of a motor response plan, we computed for each stimulus-type the proportion of the neural activity variance along the stimulus-coding direction that co-varied with the mouse's upcoming motor response. The results show that only a tiny percentage (0.5% on average) of the variations in stimulus-coding can be explained as reflecting the mouse's decision or response. Blue-shaded points denote data from individual mice. Red points are averages across mice. See also Fig. 5e. (e) Peak values of the time-dependent decoder score noise correlations (r), determined as in b, for all pairs of imaged brain areas for an example mouse, using either the data from each of five different imaging sessions, or the aggregated set of data from all imaging sessions. Fluctuations of decoder scores were correlated between sensory cortical areas during all recording sessions. The same general pattern of correlations between brain areas was visible in every session.



Extended Data Fig. 7 | Canonical correlation noise modes during the visual stimulation period for 28 different pairs of cortical areas. (a) Multiple ensembles of neurons from different cortical areas had strongly correlated noise fluctuations during visual stimulus presentation. By performing a CCA on cells' mean-subtracted activity traces for each trial-type, we identified multiple modes of significantly correlated noise modes ($P < 0.01$; comparisons of real *vs.* trial-shuffled data using the permutation test; $N = 710$ – 1340 trials) that were shared across 28 different pairs of cortical areas (abbreviated as in Fig. 1). Plots show mean \pm s.e.m. ($N = 6$ mice) correlation coefficients for the first 20 pairs of CCA noise modes for all pairs of brain areas, as determined from validation datasets that were held out from the training datasets used to identify the CCA noise modes (Methods). (b, c) In each cortical area, ~70–90% of the neurons that contributed substantially to the largest CCA noise mode were distinct from the cells that contributed to the second-largest mode. A cell was considered to contribute substantially to a CCA noise mode if its weight in the CCA mode population vector was >2 s.d. above or below the ensemble mean. (b) The mean \pm s.e.m. ($N = 6$ mice) number of cells that contributed substantially to both the first and second CCA noise modes in each brain area,

normalized by the total number of cells that contributed substantially to either of these two modes and averaged over all pairings with the other 7 brain areas. (c) Distributions of the number of simultaneously active neurons in each time bin of the stimulus presentation period for the largest five CCA noise modes shared between V1 and the other 7 cortical areas. (d) Mean correlation coefficients ($N = 6$) for neural activity in the first CCA noise mode shared between the 28 different pairs of cortical areas, for validation (*top left*) and training (*top right*) datasets, and on the set of No-Go (*bottom left*) and Go (*bottom right*) trials. The similarity of the noise correlation coefficients for all 4 subsets of trials suggests that correlated activity exists in these modes irrespective of the trial-type and that the results are not due to overfitting. (e) Highly correlated noise fluctuations between cortical areas cannot simply be explained as resulting from the activity patterns of cells on the borders between pairs of cortical areas. We repeated the analysis in (a) for all pairs of areas, while discarding the activity traces of cells in each area closer than $60 \mu\text{m}$ to the boundary of the other area identified by retinotopic mapping. The plot shows the resulting mean \pm s.e.m. ($N = 6$ mice) correlation coefficients for the CCA noise mode fluctuations between V1 and other cortical areas.

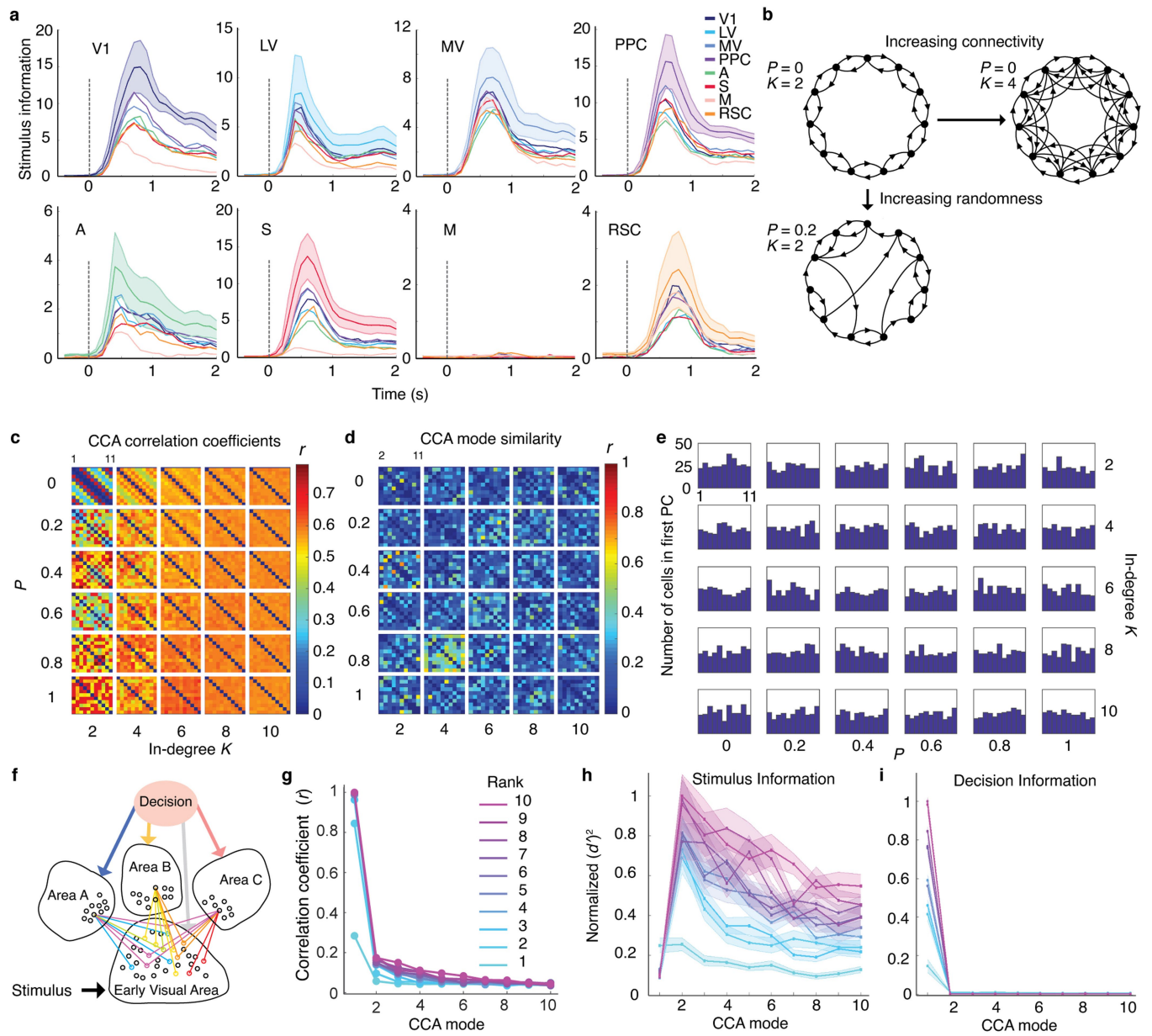


Extended Data Fig. 8 | See next page for caption.

Article

Extended Data Fig. 8 | The canonical correlation noise modes before stimulus onset were distinct from those after stimulus onset, which were task-related. (a) During the inter-trial interval (ITI), there were significantly correlated noise fluctuation modes that were shared between cortical areas. However, these modes were not the same as the shared noise fluctuations that arose at stimulus onset. The plots show the mean ($N = 6$ mice) time courses of the correlation coefficients for the first- and second-largest noise modes shared between 28 different pairs of brain areas (pairs denoted via the graph titles and the colour legend at far right), as found by applying CCA separately to ITI periods ($-2 < t < 0$) and visual stimulation periods ($2 > t > 0$). Dashed traces, with and without open circles, respectively show the correlation coefficients for the first and second shared noise modes as identified during ITI periods. Solid traces, with and without open circles, respectively show the correlation coefficients for the first and second shared noise modes as identified during stimulus periods. At stimulus onset ($t = 0$), correlated fluctuations declined within the CCA noise modes identified during ITI periods, whereas correlated fluctuations within the modes identified during the task substantially increased. (b) CCA noise modes found during stimulus periods differ from those found during ITI periods, as shown by the cross-correlation coefficients between the CCA noise modes found for each pair of brain areas before *vs.* after stimulus onset. The plots show these cross-correlation coefficient values for the largest 5 modes for each pair of brain areas. To compute these coefficients, for each mouse we created 200 different random assignments of half of the trials into a training set and half of the trials into a validation set. Using 100 of these random assignments, we determined CCA noise modes for the ITI period. Using the other 100 assignments, we determined CCA noise modes for the task period. For each entry in the plots, we plotted the mean value of the cross-correlation coefficient, averaged across all 10,000 pairings of one mode from

the ITI period and one from the stimulus period, and across 6 different mice. Within each plot, row labels designate the brain area for which we computed the cross-correlation coefficient; column labels designate the area with which the row area was paired in the CCA. (c) As a control analysis for the results of (b), we examined the variability in our estimates of the largest 5 CCA noise modes during the stimulus period. To do this, we computed for each mouse the correlation coefficients between the CCA modes determined from 100 different random assignments of trials into training and validation sets. This showed that most CCA modes are stable during the stimulus presentation period. For each entry in the plots, we plotted the mean value of the cross-correlation coefficient, averaged across all 9,900 pairings of two different mode determinations from the stimulus period, and across 6 different mice. Within each plot, row labels designate the brain area for which we computed the cross-correlation coefficient; column labels designate the area with which the row area was paired in the CCA. The results show that the relative lack of stability exhibited in (b) between CCA noise modes before versus after stimulus onset is not simply due to the statistical variability in the determination of CCA noise modes. (d) In each imaged brain area, we performed a principal component analysis (PCA) of the noise fluctuations around the mean stimulus-evoked responses, averaged over both stimuli. For each brain area, we then computed the correlation coefficients between the modes identified by PCA and those identified by CCA with each of the other 7 brain areas. The results show that fluctuation modes identified by PCA are highly distinct from those found by CCA, indicating that PCA can be incapable of detecting correlated fluctuations between brain areas. (e) Analogous plots to those in (d), except that we performed the PCA over the aggregated set of all brain areas. (f) Plots analogous to those in Fig. 5e, except that results are shown for all pairs of brain areas, rather than averaged across all pairs of sensory areas.



Extended Data Fig. 9 | See next page for caption.

Article

Extended Data Fig. 9 | Computational simulations of network dynamics show that the global CCA mode likely reflects a common signal that is broadcast to all the imaged cortical areas. (a) For the real experimental data, the graphs show the time dependence of the information, $(d')^2$, encoded about stimulus identity within CCA modes 2–10 in each brain area, plotted as a function of time relative to stimulus onset. (We omitted the first CCA mode, which does not convey stimulus information, Fig. 5d, e). To compute $(d')^2$ we trained consensus decoders based on the neural activity in each brain area during the stimulus presentation period of correctly performed trials. We then projected the neural dynamics onto each of the CCA modes and used the resulting 9-dimensional activity data to train and test instantaneous decoders of the stimulus identity. The vertical dashed lines indicate the stimulus onset. (b) To explore the patterns of interconnectivity that can give rise to a global CCA noise mode, we simulated neural activity within a range of small world networks and systematically varied the extent and randomness of the interconnections between pairs of brain areas (Methods). The schematic shows 3 example small world model networks with unidirectional connections between 11 brain areas. Each node denotes one brain area with 500 neurons. The parameter K is the 'in-degree', *i.e.* the number of projections received by each brain area. The parameter P determines the probability that the brain area sending a projection is randomly reassigned to a node outside the K nearest neighbours of the recipient brain area. The distribution of connection weights between areas was set so as to approximately match the canonical correlation coefficients observed in the real cortical recordings (Methods). A wide range of these models exhibited CCA modes among all pairs of brain areas that resembled the patterns of correlated activity fluctuations in our *in vivo* recordings of neural activity (panel c). However, no model had a global CCA mode, as each pair of brain areas generally had a unique set of co-fluctuations distinct from those in other pairs of brain areas (panel d). (c) Canonical correlation coefficients for the strongest CCA modes between all pairs of 11 areas, plotted for different values of K and P . Strongly correlated CCA fluctuations were observed between all pairs of areas in most of the simulations. (d) Correlation coefficients for the first CCA modes between one

simulated brain area and each of the other 10 brain areas, plotted as in Fig. 5a. Even when strongly correlated CCA modes exist between all pairs of areas, as shown in (c), the neural ensembles comprising these modes are largely unique and do not establish a global mode—unlike in our actual recordings (Fig. 5a) in which the first CCA mode was global and independent of the pair of brain areas chosen for CCA. These results suggest that global CCA modes may be inconsistent with information transmission through a small-world architecture. (e) The number of cells in each simulated brain area that had their first PCA weights >2 s.d. away from the mean value. Even though the simulated small world networks lacked a global CCA mode, the first mode identified by principal components analysis (PCA) was widely distributed across brain areas. Thus, the existence of distributed PCA modes does not imply the existence of a global CCA mode. (f, g) Schematic, f, of a simulated neural network (Methods) in which information about the visual stimulus is transmitted via separate channels to different higher-order cortical areas, whereas information about the sensory decision is broadcasted in parallel to these higher-order areas. The strengths of neural connections from the early visual area and each of the two higher-order areas were chosen randomly from a Gaussian distribution. The matrix of neural connections between each pair of brain areas had a rank between 1–10. g, Correlation coefficients between CCA modes in simulated cortical areas. In contrast to small-world connectivity, networks in which a single source broadcasted a common signal to multiple brain areas did have a global CCA mode, as in cortex (Fig. 5a). These results suggest the global CCA mode in cortex reflects the widespread distribution of a common signal conveying information about the mouse's upcoming response to all imaged brain areas, rather than via separate inter-area connections. (h, i) Normalized values of $(d')^2$ determined for the simulated network of (f) for distinguishing between the two different stimuli, (h), or decisions, (i), plotted for each of the 10 largest CCA modes between all pairs of areas receiving input from the Early Visual Area. Results are shown separately for networks with neural connection matrices of different ranks. Results are averaged across 25 different networks with similar architecture. Shading: s.e.m. across the 3 different simulated areas, Areas A, B and C. Fig. 5e shows similar results for the real experimental data.