Figure S1

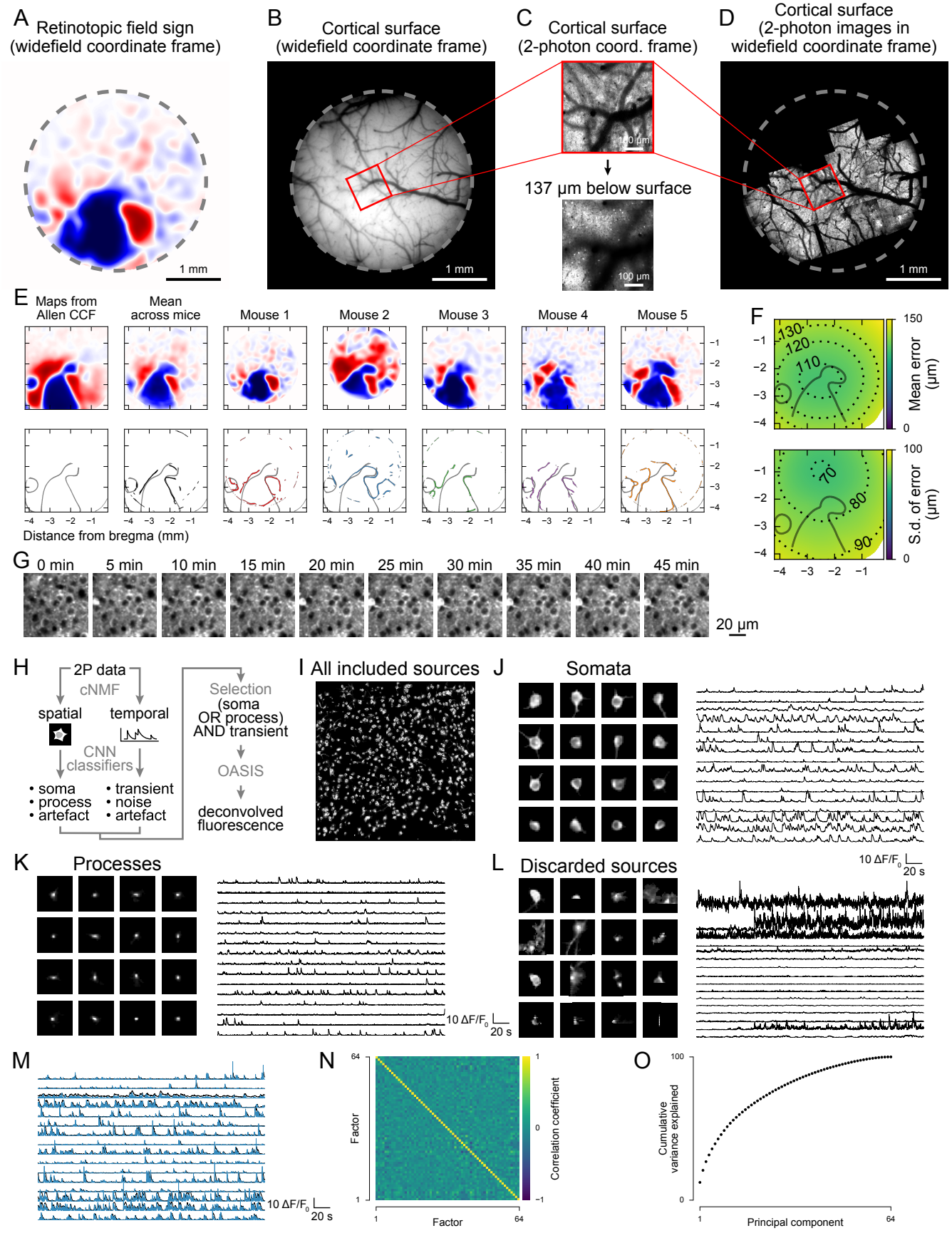

A
Retinotopic field sign
(widefield coordinate frame)
1 mm
B
Cortical surface
(widefield coordinate frame)
1 mm
C
Cortical surface
(2-photon coord. frame)
100 µm
137 µm below surface
100 µm
D
Cortical surface
(2-photon images in
widefield coordinate frame)
1 mm
E
Maps from
Allen CCF
Mean
across mice
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Distance from bregma (mm)
F
Mean error
(µm)
S.d. of error
(µm)
G
0 min
5 min
10 min
15 min
20 min
25 min
30 min
35 min
40 min
45 min
20 µm
H
2P data
cNMF
spatial
temporal
CNN
classifiers
• soma
• process
• artefact
• transient
• noise
• artefact
Selection
(soma
OR process)
AND transient
OASIS
deconvolved
fluorescence
I
All included sources
J
Somata
K
Processes
10 ΔF/F0
20 s
L
Discarded sources
10 ΔF/F0
20 s
M
10 ΔF/F0
20 s
N
Factor
Factor
Correlation coefficient
O
Cumulative
variance explained
Principal component

Figure S1. Related to Figure 2. Image alignment, source extraction and factor correlations.

(A) Example retinotopic field sign for one mouse, recorded through the cranial window using widefield fluorescence imaging.

(B) Same field of view as (A), but showing surface vessel pattern.

(C) Top: Example two-photon image of the surface vasculature. Bottom: Corresponding mean image at the depth where neural activity was recorded.

(D) Composite of all two-photon surface images for one mouse, aligned into the widefield coordinate frame. For alignment, matching features in the vasculature were determined manually in the widefield image (B) and the two-photon image (C) and used to compute a similarity transformation (translation, rotation and scaling) that registered the two-photon image into the widefield coordinate frame.

(E) Retinotopic field sign maps and outlines used to align data from each mouse into the Allen Common Coordinate Framework (CCF). Leftmost column: field sign map (top) and outlines (bottom) obtained from the Allen Institute, pre-aligned into the CCF. Other columns: Data from our mice, aligned manually (translation and rotation) to the CCF.

(F) Estimated error of the procedure used in (E). Top, mean; bottom, s.d.; n = 79 field sign maps obtained from the Allen Institute.

(G) Image timecourse for an example two-photon session. A sub-region from the center of the frame is shown. Each image is the mean of 1000 imaging frames, averaged after motion correction. Each image was independently normalized. No photodamage is apparent after 45 min of continuous imaging.

(H) Source extraction workflow.

(I) All fluorescence sources that passed the selection criteria in the session shown in Figure 2A.

(J) Example source footprints and traces. Left: Somata that passed the criteria. Right: Corresponding ΔF/F-traces.

(K) Point-shaped processes (putative apical dendrites) that passed the criteria.

(L) Sources that did not pass the criteria.

(M) Example ΔF/F-traces (black) and deconvolved traces (blue) for sources that passed the criteria.

(N) Correlation matrix for all behavioral factors.

(O) Cumulative percentage of variance explained by the principal components of the factors.

Figure S2

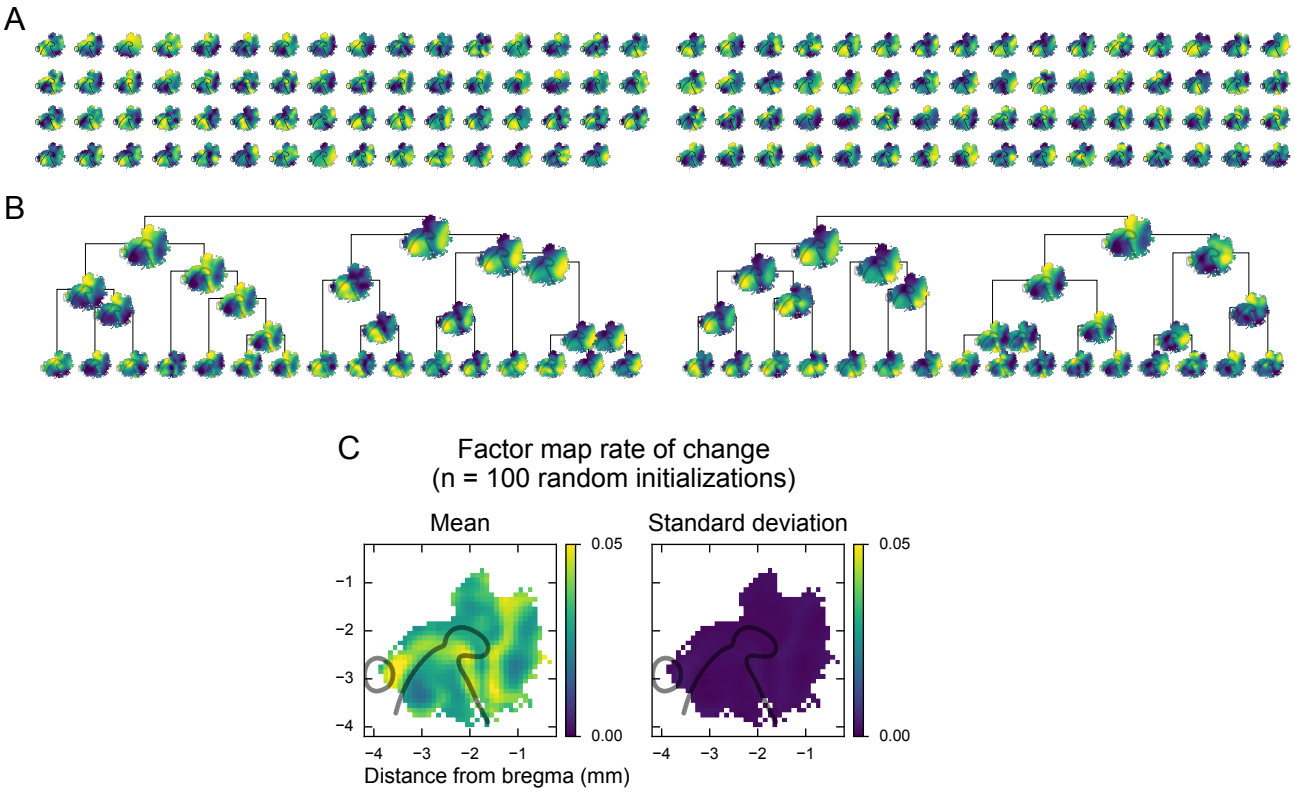

A
B
C
Factor map rate of change
(n = 100 random initializations)
Mean
Standard deviation
0.05
0.00
−1
−2
−3
−4
Distance from bregma (mm)
D
Variability in boundary location
2 areas
3 areas
4 areas
5 areas
6 areas
7 areas
8 areas
9 areas
10 areas
11 areas
12 areas
13 areas
14 areas
15 areas
0
1
Boundary probability
across n = 100 random initializations
E
Cluster discriminability
(n = 100 random initializations)
Fraction classified correctly
Data
Chance
Number of areas
2
15

Figure S2. Related to Figure 3. Analyses of factor weights are robust to non-identifiability.

The factors obtained from the encoding model (Figure 2) are not uniquely identifiable: some linearly transformed versions of the readout weight vectors provide equally good model fits (similar to non-identifiability in classical factor analysis). Therefore, re-fitting the model with different weight initializations can lead to factors that appear different when considering individual factors. However, linear transformations of the factor space do not change the intrinsic structure (relative positions) of the data points in this space. Our main analyses capture properties of this structure that are robust to the non-identifiability. Here, we provide control analyses to show this explicitly.

(A) Sets of factor maps for two example model initializations. Maps appear different for different initializations and are therefore hard to interpret directly.

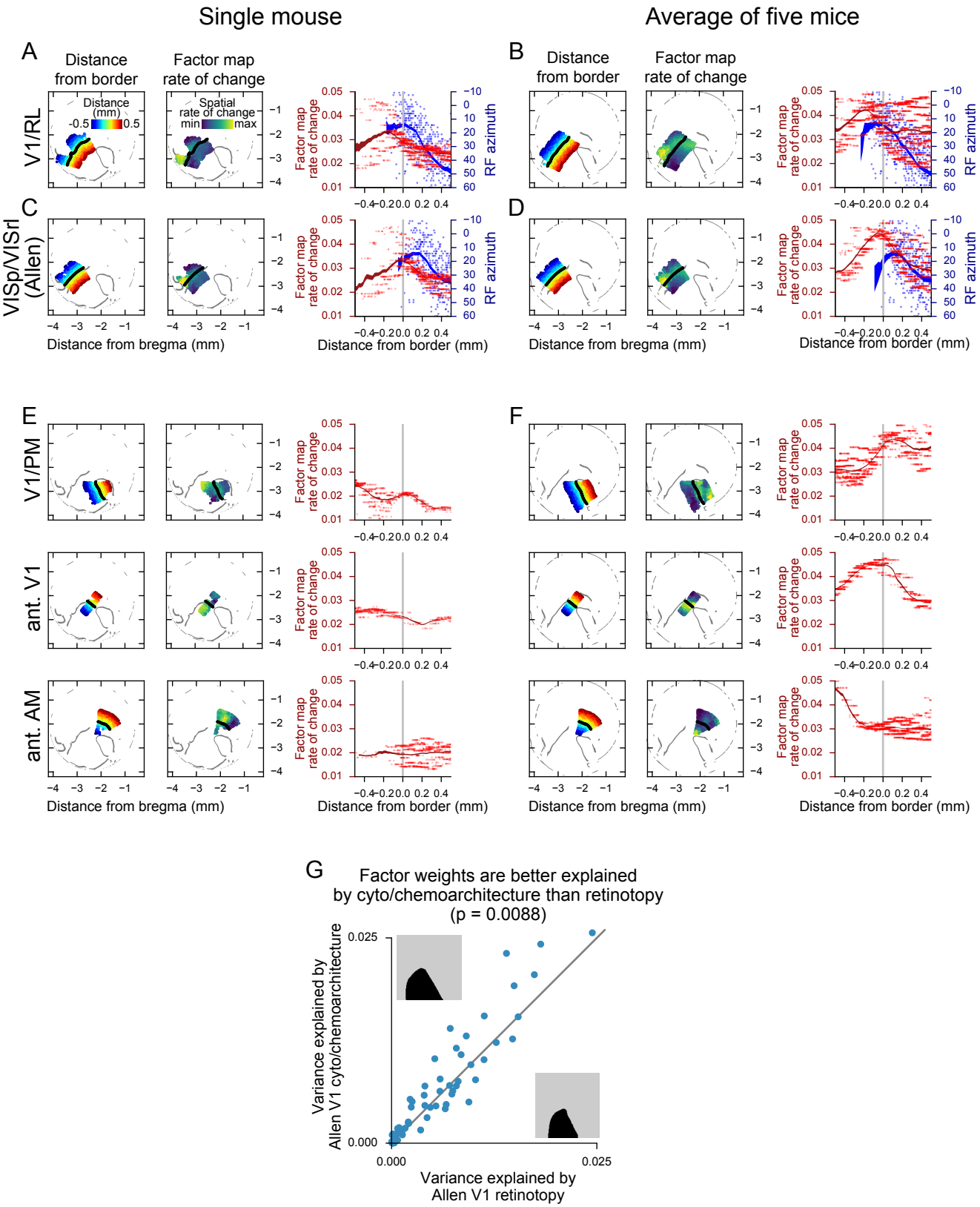
(B) Hierarchical clustering of the two sets of maps in (A). Even though individual maps in (A) vary across model initializations, they represent the same underlying encoding structure. Therefore, the patterns at the top of the clustering hierarchy are highly similar across initializations. Note that the sign of the maps is arbitrary.

(C) Related to Figure 3A: The factor map rate of change analysis is stable across model initializations. The model was re-fit 100 times with random weight initializations and the factor rate of change maps were computed as in Figure 3A. Left, mean, right, s.d. The variability across model initializations (s.d.) is much smaller than the amplitude of the features in the mean map.

(D) Related to Figure 3B: Variability in boundary locations across model initializations for the k-means clustering of factor maps. A pixel is defined to be at a boundary if at least one of its neighbors belongs to a different cluster. The plot shows the probability of pixels being at a boundary across 100 model initializations.

(E) Related to Figure 3E: Variability in classification performance. Dots indicate the median, error bars the minimum and maximum across 100 model initializations.

Figure S3


Single mouse
Average of five mice
A
B
C
D
E
F
Distance from border
Factor map rate of change
V1/RL
VISp/VISrl (Allen)
V1/PM
ant. V1
ant. AM
Distance (mm)
-0.5
0.5
Spatial rate of change
min
max
Factor map rate of change
RF azimuth
Distance from bregma (mm)
Distance from border (mm)

G
Factor weights are better explained by cyto/chemoarchitecture than retinotopy (p = 0.0088)
Variance explained by Allen V1 cyto/chemoarchitecture
Variance explained by Allen V1 retinotopy
0.000
0.025

Figure S3. Related to Figure 3. Peaks in the rate of change of factor maps are not aligned with retinotopic borders.

All data comes from the five mice used for the main figures. For the plots in A-F, colored dots show neurons within 500 μm of the specified border. In the left plot, neurons are colored by distance from the border; the border is indicated in black. In the middle plot, neurons are colored by the local rate of change of the factor maps (as in Figure 3A). In the right plot, the x-axis shows distance from the border (as illustrated by the color in the left plot), the left y-axis (red dots) shows the rate of change of the factors at the location of each neuron, and the right y-axis (blue dots) show the single-neuron receptive field (RF) azimuth (where available). Line width indicates mean ± s.e.m.

(A) V1/RL border, as defined by widefield retinotopic mapping, in a single mouse. The border is defined as the line where the field sign magnitude is closest to zero. The peak of the rate of change of the factor maps (red line in the plot on the right) is displaced laterally by approximately 200 μm from the retinotopic field sign border (gray line), as reported for cytoarchitectonic and chemoarchitectonic borders of V1 (Zhuang et al., 2017). In contrast, the RF azimuth measured for single neurons by two-photon imaging (blue line) peaks at zero, indicating accurate alignment between two-photon sessions and widefield retinotopy data.

(B) As in (A), but for with data from five mice.

(C) VISp/VISrl border, as defined in the Allen CCF (Allen Institute, 2017). This border was defined by the Allen Institute based on cytoarchitectonic and chemoarchitectonic data. In contrast to (A), the peak of the rate of change of the factor map (red line in the plot on the right) aligns with this border, whereas the peak in RF azimuth (blue line) is displaced medially.

(D) As in (C), but for data from five mice.

(E) Rate of change of the factor maps at further retinotopic borders, for a single mouse.

(F) Rate of change of the factor maps at further retinotopic borders, for five mice. Peaks in the rate of change of the factor maps follow a pattern that is more similar to cytoarchitectonic and chemoarchitectonic features as described previously (Allen Institute, 2017; Zhuang et al., 2017) than to retinotopic field sign borders. For example, there is a peak in the rate of change of the factor maps at the anterior (ant.) border of V1, but none at the anterior border of AM.

(G) Factor weights are better explained by cyto/chemoarchitecture than retinotopy. Each dot represents one of the 64 task factors. For each factor, the squared correlation (variance explained) was computed between the true factor weights for all neurons and two simple templates: The first template (plotted on the x-axis) was based on the retinotopic outline of V1 and modeled neurons within V1 with -1 and neurons outside of V1 with +1. The second template (y-axis) was similar, but based on the outlines of VISp as defined in the Allen Institute CCF, which are based on cyto/chemoarchitectonic features. The Allen Institute template explained more variance in the factor weights ($p = 0.0088$, two-tailed t-test). This analysis is complementary to the analyses of the factor map rate of change above because it uses raw single-neuron factor weights without smoothing or rate of change computation.

Figure S4

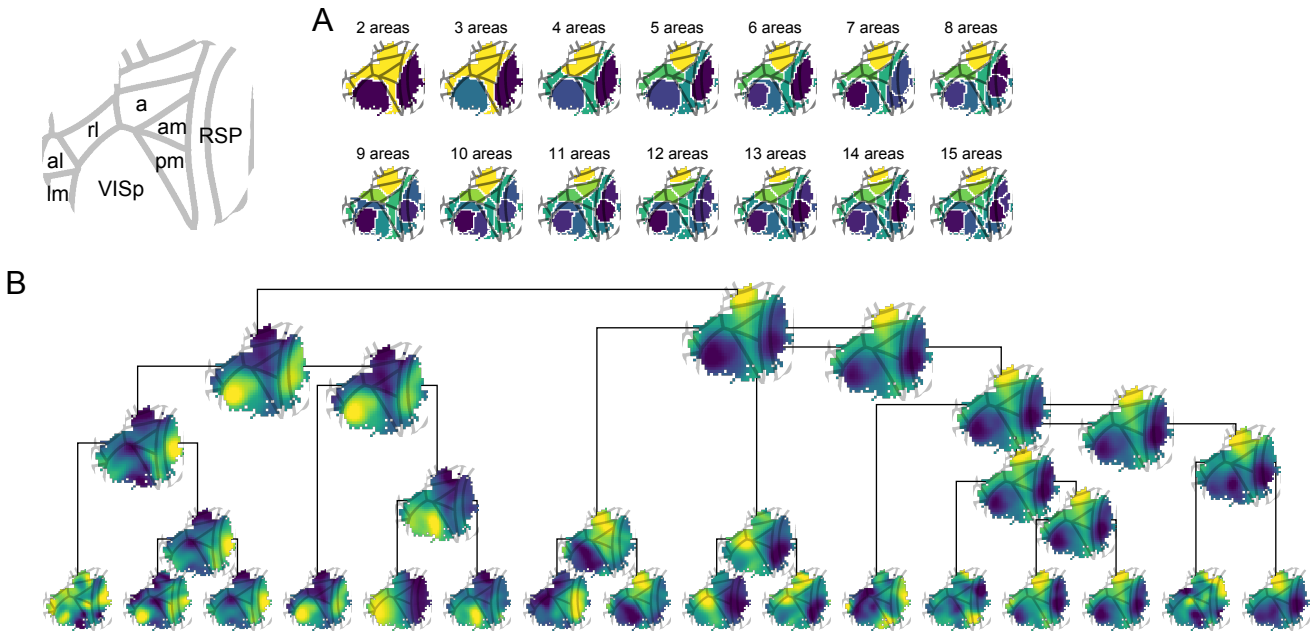

A
2 areas
3 areas
4 areas
5 areas
6 areas
7 areas
8 areas
9 areas
10 areas
11 areas
12 areas
13 areas
14 areas
15 areas
a
rl
am
RSP
al
pm
lm
VISp
B

Figure S4. Related to Figure 3. k-means and hierarchical clustering with Allen Institute CCF outlines.

(A) Same k-means clustering of factor maps as in Figure 3B, but with Allen Institute CCF areas overlaid (gray lines; see legend on the left).

(B) Same agglomerative hierarchical clustering of the 64 factors as in Figure 3F, but with Allen Institute CCF areas overlaid.

Figure S5

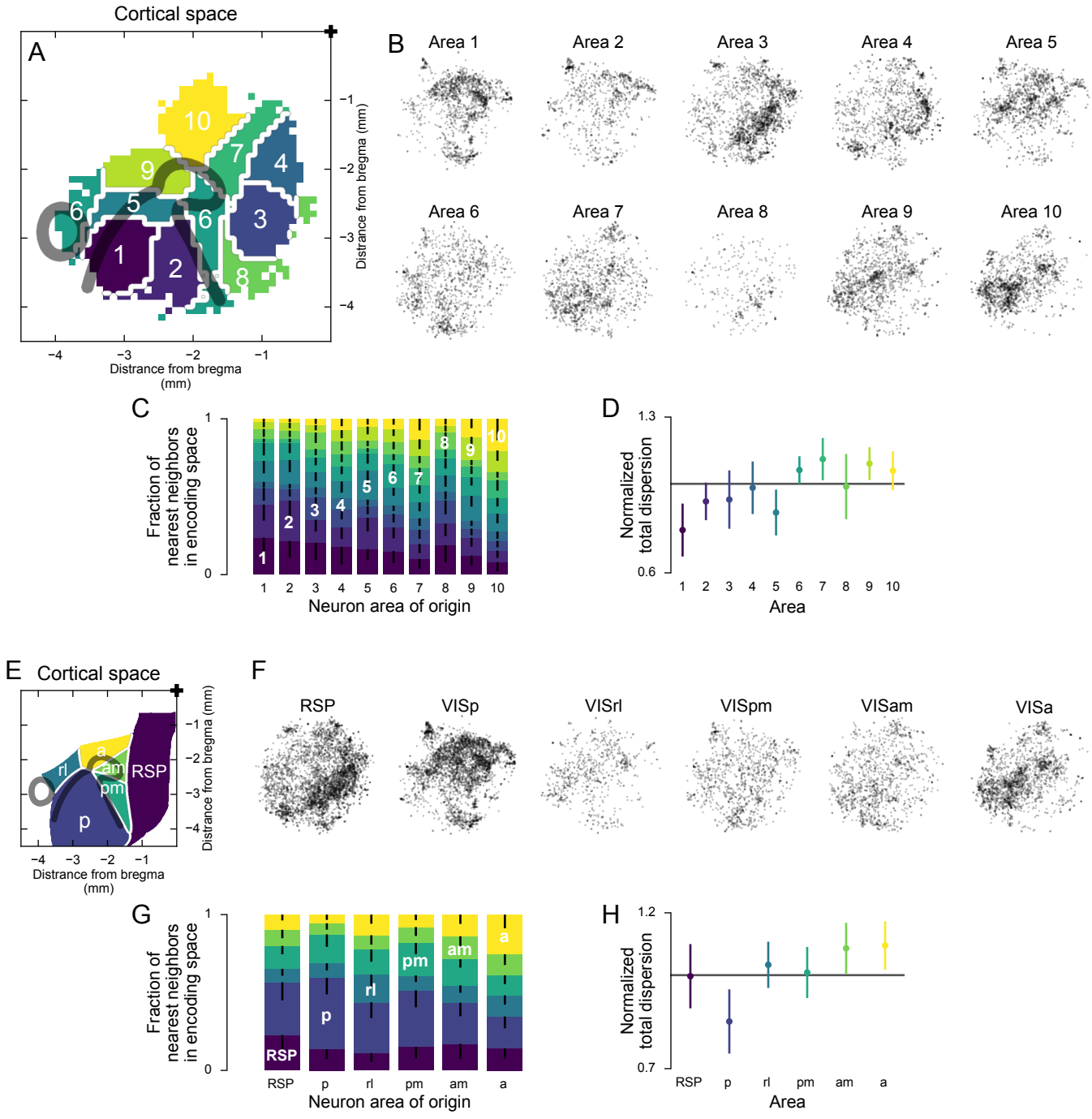

A
Cortical space
Distrance from bregma (mm)
Distrance from bregma (mm)
B
Area 1
Area 2
Area 3
Area 4
Area 5
Area 6
Area 7
Area 8
Area 9
Area 10
C
Fraction of nearest neighbors in encoding space
Neuron area of origin
D
Normalized total dispersion
Area
E
Cortical space
Distrance from bregma (mm)
Distrance from bregma (mm)
F
RSP
VISp
VISrl
VISpm
VISam
VISa
G
Fraction of nearest neighbors in encoding space
Neuron area of origin
H
Normalized total dispersion
Area

Figure S5. Related to Figure 4. Encoding property distributions for 10 areas and for Allen Institute parcellation.

(A) 10-area parcellation, as in Figure 3B.

(B) t-SNE embedding of the factor weights as in Figure 4C, but for 10 areas.

(C) Anatomical location of nearest-encoding-space-neighbors as in Figure 4E, but for 10 areas.

(D) Normalized total dispersion of weights as in Figure 4F, but for 10 areas. Based on hierarchical bootstrap with 1000 samples, the mean is different from 1 with: area 1, $p < 10^{-3}$, n = 1559; area 2, p = 0.071, n = 893; area 3, p = 0.242, n = 2071; area 4, p = 0.340, n = 1500; area 5, p = 0.018, n = 1458; area 6, p = 0.030, n = 1210; area 7, p = 0.079, n = 1345; area 8, p = 0.395, n = 332; area 9, p = 0.035, n = 1402; area 10, p = 0.104, n = 1568 neurons.

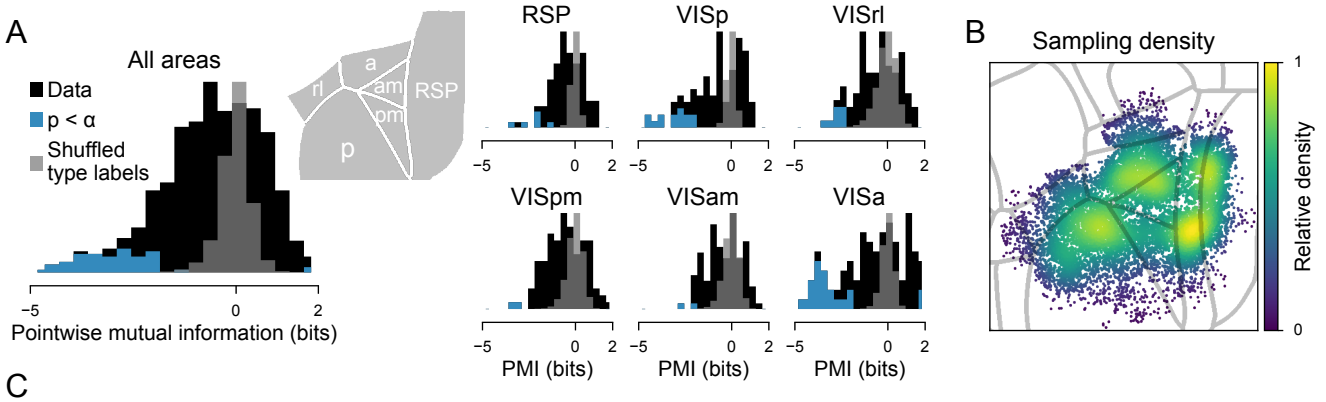
(E) Area parcellation from the Allen Institute CCF (Allen Institute, 2017).

(F) t-SNE embedding of the factor weights as in Figure 4C, but for the Allen Institute areas.

(G) Anatomical location of nearest-encoding-space-neighbors as in Figure 4E, but for the Allen Institute areas.

(H) Normalized total dispersion of weights as in Figure 4F, but for the Allen Institute areas. Based on hierarchical bootstrap with 1000 samples, the mean is different from 1 with: RSP, p = 0.432, n = 4210; VISp, p = 0.007, n = 3544; VISrl, p = 0.229, n = 697; VISpm, p = 0.455, n = 1092; VISam, p = 0.047, n = 1173; VISa, p = 0.060, n = 1913;

Figure S6


A
All areas
Data
p < α
Shuffled type labels
−5
0
2
Pointwise mutual information (bits)
rl
a
am
pm
RSP
p
RSP
VISp
VISrl
VISpm
VISam
VISa
PMI (bits)
B
Sampling density
Relative density
1
0

C

Figure S6. Related to Figure 5. Encoding type distribution histograms, using Allen Institute parcellation, and cortical distribution of all 100 encoding types.

(A) Histograms of pointwise mutual information between area and encoding type as in Figure 5C and D, but for the Allen Institute areas.

(B) Sampling density map showing all neurons displayed in the maps in (C). Overlapping dots are colored based on a Gaussian kernel density estimation (200 µm s.d.).

(C) Maps of all encoding types. To order the maps, the most significant PMI value among the five areas was found for each encoding type, and types were sorted in descending order of that value. Types are therefore ordered from strongly area-specific to strongly area-excluded. Red outlines, retinotopic boundaries in our recordings; blue outlines, area parcellation from the Allen Institute CCF.

Figure S7

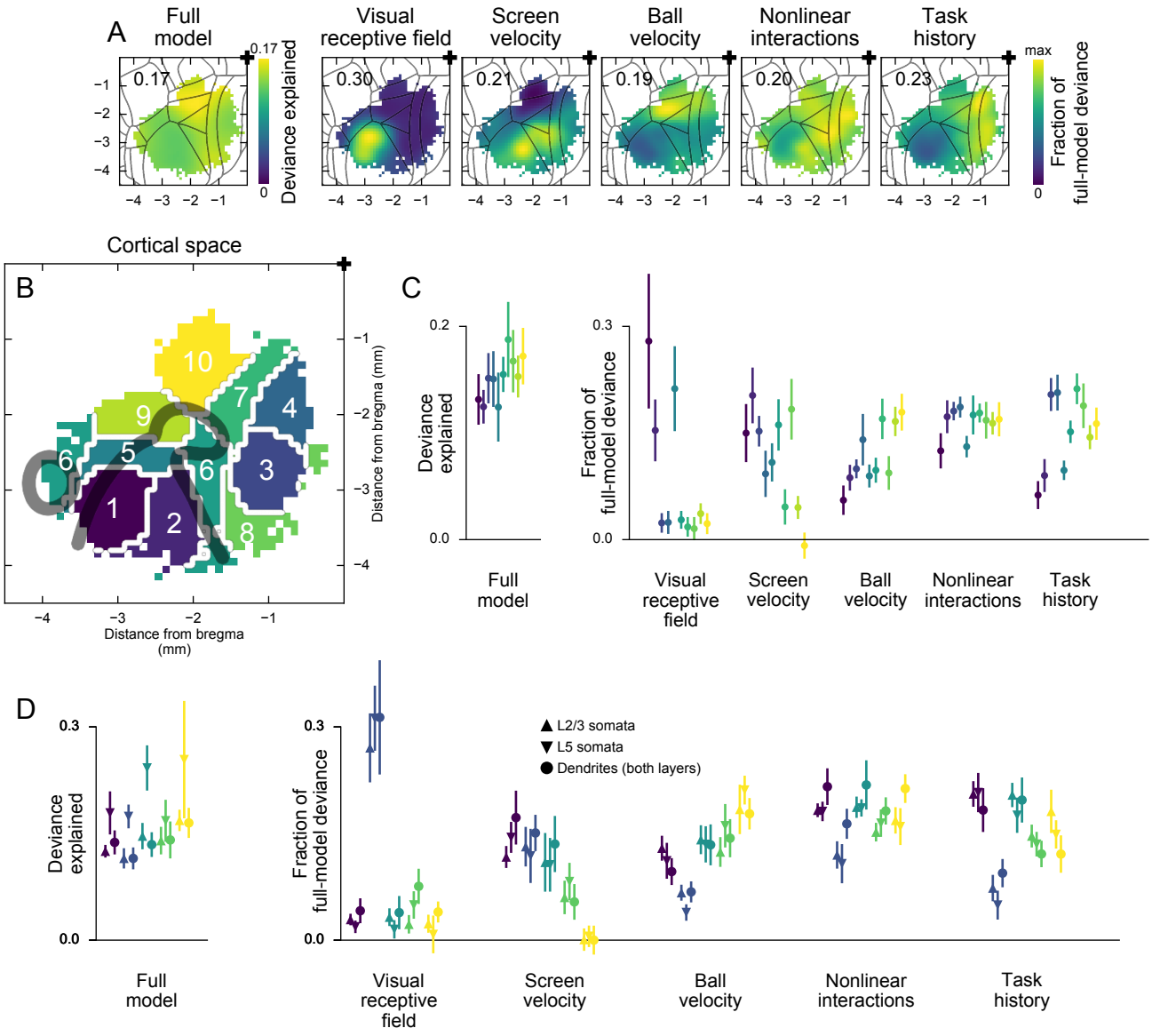

A
Full model
Visual receptive field
Screen velocity
Ball velocity
Nonlinear interactions
Task history
0.17
0.30
0.21
0.19
0.20
0.23
Deviance explained
Fraction of full-model deviance
max
B
Cortical space
Distance from bregma (mm)
C
Deviance explained
Fraction of full-model deviance
D
L2/3 somata
L5 somata
Dendrites (both layers)

Figure S7. Related to Figure 6. Encoding of task variables.

(A) Maps of encoding strength as in Figure 6B, but with Allen Institute CCF outlines overlaid.

(B) Parcellation of cortex used in (C). From Figure 3B.

(C) Data from Figure 6B, binned by areas in (B). Colors as in (B). Error bars, 5th and 95th percentile of hierarchical bootstrap of the mean. Number of neurons: area 1, 2703; area 2, 1772; area 3, 3671; area 4, 2731; area 5, 2776; area 6, 2123; area 7, 2023; area 8, 499; area 9, 2373; area 10, 2313. Mean deviance explained is greater than zero in all areas with $p < 10^{-3}$. Mean fraction of full-model deviance is greater than zero for all areas and conditions with $p < 0.011$ except for visual receptive fields, area 7 ($p = 0.021$), area 8 ($p = 0.062$) and screen velocity, area 10 ($p = 0.810$).

(D) Same as in Figure 6D, but divided by cortical layer. Since dendrites (point-shaped fluorescence sources) could have their somata in either layer (Peron et al., 2015), they are shown separately. Encoding was similar in L2/3 somata, L5 somata, and dendrites, except for the encoding of nonlinear feature interactions, which was consistently stronger in dendrites than in somata. This result could reflect nonlinear integration of synaptic inputs observed in cortical dendrites, but further investigation is needed. Upward-pointing triangles, L2/3 somata; inverse triangles, L5 somata; circles, dendrites. Number of L2/3 somata: area 1, 4175; area 2, 2762; area 3, 1888; area 4, 1334; area 5, 1407. Mean fraction of full-model deviance is greater than zero for all areas and conditions with $p < 0.005$ except for area 5, screen velocity ($p = 0.49$). Number of L5 somata: area 1, 872; area 2, 1037; area 3, 778; area 4, 718; area 5, 308. Mean fraction of full-model deviance is greater than zero for all areas and conditions with $p < 10^{-3}$ except for: area 3, visual receptive fields ($p = 0.025$); area 5, visual receptive fields ($p = 0.124$); area 5, screen velocity ($p = 0.236$). Number of dendrites: area 1, 438; area 2, 926; area 3, 323; area 4, 703; area 5, 674. Mean fraction of full-model deviance is greater than zero for all areas and conditions with $p < 10^{-3}$ except for: area 3, visual receptive fields ($p = 0.004$); area 5, screen velocity ($p = 0.510$).

Figure S8

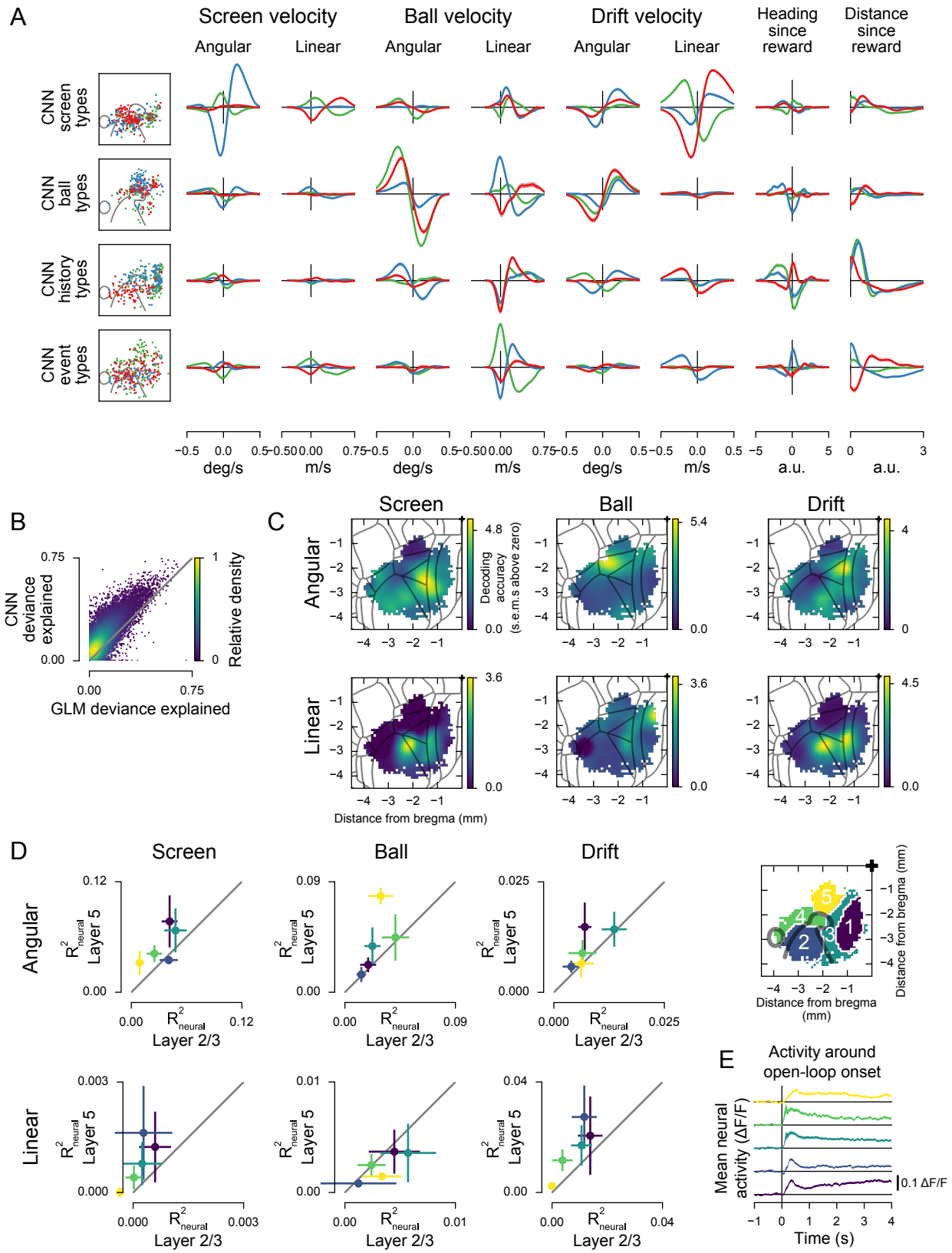

A
Screen velocity
Ball velocity
Drift velocity
Heading since reward
Distance since reward
Angular
Linear
CNN screen types
CNN ball types
CNN history types
CNN event types
deg/s
m/s
a.u.
B
CNN deviance explained
GLM deviance explained
Relative density
C
Screen
Ball
Drift
Angular
Linear
Decoding accuracy (s.e.m.s above zero)
Distance from bregma (mm)
D
Layer 5
Layer 2/3
E
Activity around open-loop onset
Mean neural activity (ΔF/F)
0.1 ΔF/F
Time (s)

Figure S8. Related to Figure 8. GLM tuning curves and decoding maps for further task variables.

(A) GLM tuning curves for the encoding types with the largest fraction of full-model deviance for each feature in the CNN-based model, as defined in Figure 6D. Left: Distribution across cortex of the neurons in each type, as in Figure 7. Right: GLM tuning curves corresponding to the types shown on the left. Each curve is the average for the neurons in that type. Shading shows s.e.m. Error bounds are small because the encoding types have similar encoding properties, by design.

(B) Comparison of the CNN and GLM fit quality. Each point represents one neuron; overlapping points are colored by density.

(C) Maps of decoding accuracy as in Figure 8B, but with Allen Institute CCF outlines overlaid.

(D) Comparison of decoding performance (R^2_{neural}) across layers. Error bars, 5th and 95th percentile of hierarchical bootstrap of the mean. Colors indicate areas as shown in the legend on the right. Number of neurons: L2/3: area 1, 2067; area 2, 1493; area 3, 915; area 4, 696; area 5, 856. L5: area 1, 599; area 2, 799; area 3, 653; area 4, 436; area 5, 250.

(E) Mean activity traces, aligned to open-loop onsets and split by area. Colors as in the map above. Shading, mean ± s.e.m. across neurons.