Neuron, Volume *107*

Supplemental Information

Cortical Observation by Synchronous Multifocal Optical Sampling Reveals Widespread Population Encoding of Actions

Isaac V. Kauvar, Timothy A. Machado, Elle Yuen, John Kochalka, Minseung Choi, William E. Allen, Gordon Wetzstein, and Karl Deisseroth

Supplementary figures for: Cortical observation by synchronous multifocal optical sampling reveals widespread population encoding of actions

Isaac V. Kauvar[1,2,*], Timothy A. Machado[1,*],
E. Yuen[1], J. Kochalka[1,3], M. Choi[3,4], W.E. Allen[1,3,4], Gordon Wetzstein[2], Karl Deisseroth[1,5]

[*] equal contributions

[1] *Department of Bioengineering, Stanford University*
[2] *Department of Electrical Engineering, Stanford University*
[3] *Neuroscience Graduate Program, Stanford University*
[4] *Department of Biology, Stanford University*
[5] *Department of Psychiatry, Stanford University and Howard Hughes Medical Institute*

Corresponding Author:
Karl Deisseroth (deissero@stanford.edu)

Figure Legends

Figure S1. **Surgical procedure for implantation of large-scale window (relates to Figure 1).**
(A) Geometry of curved glass window implant. (B) Screenshot of software interface for selecting keypoints. (C) Coordinates of keypoints in Panel B, with approximate thickness of the skull at each keypoint. Thickness varies for mice of different ages and sizes. (D) Photograph of robotic stereotax and setup. (E) After cutting skin back, ear bars are placed above skin and muscle, and skull is cleaned. (F) Conductive clip is attached to inside of the skin, and keypoints are drilled using robotic stereotax. Depth is determined using a conductivity-based autostop mechanism, or alternatively, inspection and knowledge of the rough depth of each keypoint based on previous mice. (G) Based on keypoints, robotic stereotax interpolates and automatically cuts full craniotomy. (H) After skull flap is loose, it is pulled off quickly using forceps and blood is cleaned using absorptive sponge swabs. (I) Curved glass implant is positioned using a blunt needle connected to vacuum, and pushed down such that glass is in contact with brain across the full window. (J) Window is cemented into position. (K) A depth-colored 120μm thick two-photon stack demonstrating density and health of the neurons. (L) High magnification histological section demonstrating GCaMP expression, with DAPI nuclear staining.

Figure S2. **Theory, simulation, and design underlying the enhanced performance of COSMOS (relates to Figure 1).**
(A) Median background per pixel (across recovered neurons) for three recording epochs per detection design, from an in-focus ROI. Background is estimated using the CNMF-E algorithm. (B) Total signal per neuron (across recovered neurons) for three recording epochs per detection design, from an in-focus ROI. Signal is derived based on the spatial footprint and raw temporal traces estimated using the CNMF-E algorithm. (C) Simulated comparison of light collection efficiency of a conventional macroscope with different aperture settings, a tunable lens with a 16mm aperture, a multifocal two-camera beamsplitter design with f/2 objective aperture setting, and the multifocal dual lenslet design of COSMOS. Although the beamsplitter has a depth of field of f/2 it only has half the light collection. (D) Simulated defocus blur vs. lateral distance along the curved glass window (with 11 mm radius of curvature) for each design. (E) Simulated median (taken laterally across the window) of photon density for each design. (F) Simulated signal to noise ratio vs. lateral position for each design. Assumes a Gaussian point spread function. Secondary peaks are visible for the two multifocal designs. Although f/1.4 has a higher peak SNR when in focus, SNR drops off quickly as the image defocuses. (G) Simulated median (taken laterally across the window) of the SNR for each pixel. Lenslet outperforms the other designs. (H) CAD model of dual-lenslet system. (I) Specifications of an individual lenslet. 7.09 mm from the edge of an off-the-shelf aspherical achromatic lenslet was milled away. (J) Schematic and dimensions of dual-lenslet mount. (K) CAD model of light-blocking cone used to mitigate imaging artifacts induced by visual stimuli. (L) CAD model of headbar used for head-fixing the mouse. (M) Photograph of COSMOS microscope and behavioral setup. (N) Raw sensor image of USAF 1950 Resolution test chart.

Figure S3. **Intrinsic imaging and atlas alignment method (relates to Figure 1).**
(A) Experimental setup for intrinsic imaging with a moving bar visual stimulus. (B) Flowchart of intrinsic imaging-based alignment procedure. (C) Example raw phase map with midline and V1/PM border overlaid, also indicated on the atlas. (D) (top) Alignment maps for all mice. (middle) The 700 nm illuminated images used for intrinsic imaging which were aligned to (bottom) the COSMOS fluorescence images. (E) Example raw traces with overlaid smoothed traces (both as output from CNMF-E algorithm) of a few example sources for different imaging conditions. The f/8 macroscope setting shows a clear degradation of signal-to-noise ratio in the raw traces.

Figure S4. **Additional analysis of visual-stimulus source characterization (relates to Figure 2).**
(A) Simulation-based estimate of number of neurons per source. First column: Histograms of orientation selectivity index (OSI) for two-photon data shown in Figure 2F (bottom row) generated by resampling with replacement 500 times. Second column: on each of 500 iterations, groups of neurons (quantities chosen by sampling uniformly from the range indicated on the subplot) were averaged together. These 500 mixed sources then used to generate OSI histograms. Third column: true OSI histograms derived from COSMOS data. Red percentages denote the fraction of neurons with OSI > 0.8. Fourth column: similarity between simulated COSMOS OSI distributions and the true COSMOS distribution measured by KL Divergence (lower is more similar). The most similar range of sources to mix together is denoted by a red dot. White dots denote ranges that are statistically indistinguishable from the red dot range (one-sample T-test vs. model marked with red dot, $p > 0.05$). 10 simulated histograms were generated for each point in parameter space. Results shown for three mice. (B) Format matches the above panels using data from Mouse D but where the target COSMOS distribution has all sources with OSI > 0.8 removed. The best fitting simulated histogram corresponds to one with a larger degree of neuron mixing. (C) Atlas aligned to cleared skull of a Cux2-CreER; Ai148 (GCaMP6f) mouse. Highlighted sources were identified as visually responsive (one-way ANOVA, $p < 0.01$) to a drifting grating stimulus presented in eight directions. The orientation selectivity of each visually-responsive source is indicated in color. (D) Stimulus-averaged fluorescence responses to each grating direction from sources extracted from right visual cortex of the mouse from Panel C under the COSMOS microscope. Each source is normalized to its peak response. Error

bars represent S.E.M. (E-F) Format matches Panels C-D but for a Thy1 (GCaMP6s) mouse imaged through its cleared skull. (G) Trial-averaged, max-normalized, fluorescence traces taken from visually responsive neurons found in right visual cortex of Mouse 1 (Cux2-CreER; Ai148, curved window) under a two-photon microscope. The right panel shows neural population responses for trial-averaged data shown in the left panel projected onto the basis vectors defined by PCA. (H) The same visual stimuli were presented to Mouse 1, under the COSMOS microscope, to reveal the trial-averaged, max-normalized fluorescence traces of sources taken from both visual cortical areas of Mouse 1 (left) and corresponding trajectories in PC space (right).

Figure S5. **Further details about task-responsive sources and their spatial distributions (relates to Figure 3).**
(A) Four additional mice, same format as Figure 3J, showing average traces of sources assigned to each of five task-related classes. This includes one Rasgrf2-2A-dCre:ai93, a different genetic line that labels a potentially non-identical subset of superficial neurons as Cux2-CreER. We see similar response profiles in both genotypes. We did not do any further analyses with the Rasgrf mouse data. (B) Spatial distribution of sources assigned to each task-related class. (C) Spatial autocorrelation analysis of distributions shown in Panel B (CDF is cumulative fraction of source-pairs separated by a given distance). Only five task classes had any significant deviations from the null distribution across all mice and distances (corrected $p < 0.05$, permutation test, significant values are indicated with black line overlay), and the significant task classes or CDF values were not consistent across mice. (D) Spatial autocorrelation analysis on simulated data to investigate sensitivity to clusters of different sizes. Clusters at least 1mm in diameter (1/3 of sources within the cluster area were assigned to the cluster) displayed clear significant ($p < 0.05$) differences from the null distribution (marked with black line overlay). (E) Distribution of task classes within each region (M=motor, P=parietal, R=retrosplenial, S=somatosensory, V=visual). Error bars are S.E.M. across n=4 Cux2-CreER mice. (F) Additional spatial distributions and examples of sources exhibiting “lick OFF” dynamics.

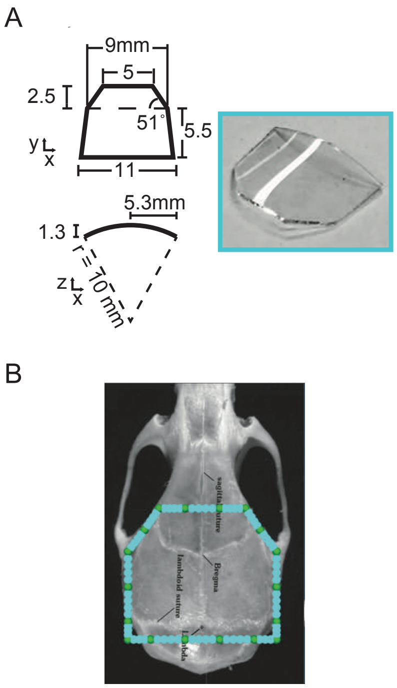
Figure S6. **Additional data related to unaveraged correlation structure (relates to Figure 4).**
(A) Trial-averaged activity correlations: (top) spatial distribution, (bottom) correlation vs. distance to the seed. Seed location is represented as a black dot in right visual cortex, referenced by the labelled arrow. (B) Unaveraged activity correlations: (top) spatial distribution), (bottom) correlation vs. distance to the seed. This source exhibits symmetrically bilateral localized correlation. (C) Example of unaveraged correlated activity between the seed source and a neighboring source (locations indicated on atlas inset). Z-scored fluorescence and inferred spikes are shown. Red arrows indicate timepoints when the seed source and its neighbor are active simultaneously. (D) Single trial responses of seed source. (E-H) Same, but for an additional mouse, with seed source in retrosplenial cortex. (I-L) Same, but for an additional mouse, with seed source in motor cortex.

Figure S7. **Additional analyses related to decoding motor actions and plans (relates to Figures 5-7).**
(A) Ability of each neural source to discriminate between any of four behavioral states (licks towards one of the three spouts, or no lick). P-value computed for each source using Kruskal-Wallis H-test and FDR-corrected. Dashed horizontal line represents $p = 0.05$. (B) Spatial distribution of sources with significant motor action discrimination capacity ($p < 0.05$). (C) Example decoding of lick direction from neural activity at 30 Hz (on test dataset). Vertical lines represent trial onsets. (D) Model prediction accuracy versus the number of PLS basis vectors used. Five models were trained for each basis set size. (E) Pre-odor neural decoding performance as a function of number of neuronal sources used. Red lines indicate means across mice. Gray lines indicate performance of models trained and evaluated against circularly permuted true-spout labels. Black lines similarly represent models trained against randomly shuffled true-spout labels. Red lines indicate means across mice. Error bars show 99% bootstrapped confidence intervals. Vertical line at 500 indicates that all analyses that pooled across cortical regions used only the top 500 sources with most discrimination capacity for decoding to avoid model overfitting. (F) Spatial distribution of sources with significant discrimination capacity ($p < 0.05$) for the spout that was most-licked during the pre-reward period. P-values were computed for each source using a Kruskal-Wallis H-test and FDR-corrected. (G) Comparison of decoding performance of either the active spout (left) or the most-licked spout (right). Corrected paired t-test p-value shown between the active and preferred conditions. Note that trials where the active spout and most-licked spout were not identical were never used for the training of any PLS models. (H) Neural decoding of the preferred spout as a function of fraction of pre-reward licks towards the preferred spout. (I) Model AUC when evaluated on "correct go trials" (defined as those where at least 70% of total licks were to the active spout), "no go" trials, "2nd trials", and "incorrect go trials" ("go" trials where less than 70% total trial licks were towards the active spout). "Incorrect go trials" are defined differently here, versus the main text (where less than 30% was the criterion), because multiple unique trials towards each spout were needed to accurately compute the AUC. (J) The total variance explained, shown for the top four basis vectors computed using either PLS (bottom, gray) or Principal Components Analysis (PCA, top, black), averaged across all four mice.

Figure S8. **Active spout decoding from high speed videos of mouse behavior and cross-temporal neural decoding (relates to Figures 6 and 7).**
(A) Extracting motion-related features from high-speed behavior videos. (B) Variance explained by behavioral video motion

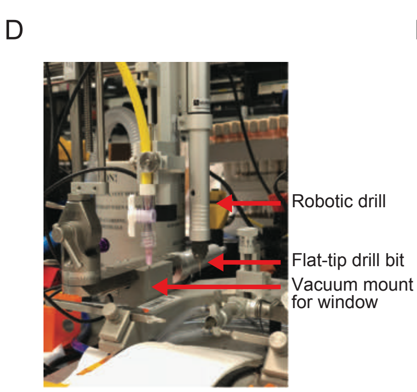
energy principal components. (C) Results from PLS decoding of active spout from pre-odor video motion features. (D) Results of decoding using pre-odor motion features from specific regions of interest. (E) We asked if there might be a difference between widespread neural representations of motor plans, and of motor plan execution. To do this, we compared sets of PLS models trained on different epochs of the trial. By training models on a single time bin (e.g. during the pre-odor period) and evaluating prediction performance on a different time bin (e.g. during movement), we tested the similarity of the PLS models over the trial. The goal was to understand whether a single projection of the data might best discriminate the trial types, or whether cortex might represent motor plans during the intertrial interval differently from actual directed licking during the peri-movement period (between lick onset and lick offset). Examining how well each of the time-specific models predicted other time bins revealed clear structure. In grouping these time bins into pre-, peri-, and post-task epochs based on the average time of lick onset and offset, we found a dissociation between models trained on intertrial interval data (from the pre- and post-task epochs) versus those trained using peri-task period data. Specifically in this panel, AUC was computed for PLS bases trained to predict the active spout position from neural data in one time window, and then tested on neural data from all time windows and each mouse. AUC is normalized by the maximum AUC for each testing time window, and averaged across four mice. Dashed lines represent the average time of lick onset and offset within the trial. For this analysis (panels E-G), testing was only done on trials where the active spout and preferred spout were identical (trials where the active spout and preferred spout were not the same were always excluded from training in all analyses). (F) Quantification of results in Panel E for individual mice, comparing cross-temporal prediction performance of three distinct time blocks: pre-task, peri-task, and post-task. For each testing time block, we statistically compared the mean prediction performance of bases trained on each time block (corrected t-test). These results are consistent with the interpretation that motor plans and consequent movements are represented by two distinct, but related, neural population representations. (G) Unnormalized cross-temporal AUC for each mouse. ns denotes corrected $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Figure S9. **Inhibition of non-motor areas impairs lick-to-target task performance (relates to Figures 5-7).**
(A) In addition to recording neuronal activity, we tested if the task-activity we found distributed across many non-motor areas was causally involved in task performance. We investigated this because it is known that specific areas of cortex can modulate the firing of neurons in primary and secondary motor cortex. For example, projections from S1 can drive M1 activity and drive the initiation of whisking (Sreenivasan et al., 2016). But in addition, posterior regions like retrosplenial cortex and primary sensory areas (Barthas and Kwan, 2017; Yamawaki, Radulovic, and Shepherd, 2016; Zingg et al., 2014), also project to secondary motor cortex and may thus play an important role in producing any dynamics observed in motor areas. We therefore set out to study whether mice could accomplish our task without normal input from all of these non-motor areas by simultaneously inhibiting multiple cortical regions using the Digital Micromirror Device (DMD) system shown here. This approach avoided the possibility that uninhibited areas in cortex might compensate for the acute shutdown of a single area. (B) Atlas alignment. (C) Intensity calibration. (D) Stimulation protocol. (E) Stimulation patterns. (F) Success rate on peri-odor stimulation vs. no stimulation trials, for each stimulation pattern, n=5 mice, corrected paired t-test between stimulation patterns. (G) Success rate on pre-odor stimulation vs. no stimulation trials, for each stimulation pattern, n=5 mice, corrected paired t-test between stimulation patterns. (H) Success rate on peri- and pre-odor stimulation for a mouse not expressing ChR2. (I-L) Summary of lick selectivity averaged across n=5 mice (with S.E.M. shown) after odor delivery, but before (top) or after (bottom) reward delivery during non-stimulation trials or stimulation trials, for each stimulation pattern. Colors denote the normalized number of licks towards each spout on trials when a given spout is active. (M-O) Example lick rasters demonstrating that licking is not fully abolished during photoinhibition, but rather the specificity of it is disrupted (for each stimulation pattern, some sessions showed incorrect preference for a single spout, while others showed nonspecific licking towards multiple spouts). (P) Comparison of total number of peri-odor and post-reward licks during no-stimulation (gray) and stimulation (blue) trials (mean +/- S.E.M. across mice, corrected paired t-test between stimulation patterns). ns denotes corrected $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.



C

Y (mm)	X (mm)	Depth (μm)
2.70	-2.50	200-400
1.45	-3.35	450-700
0.20	-4.30	400-700
-1.85	-4.30	400-600
-3.85	-4.30	250-600
-4.90	-4.30	800-1000
-4.90	-3.00	500-800
-4.90	-1.00	500-800
-4.90	1.00	500-800
-4.90	3.00	500-800
-4.90	4.30	800-1000
-3.85	4.30	250-600
-1.85	4.30	400-600
0.20	4.30	400-700
1.45	3.35	450-700
2.70	2.50	200-400
2.70	1.00	100-200
2.70	-1.00	100-200



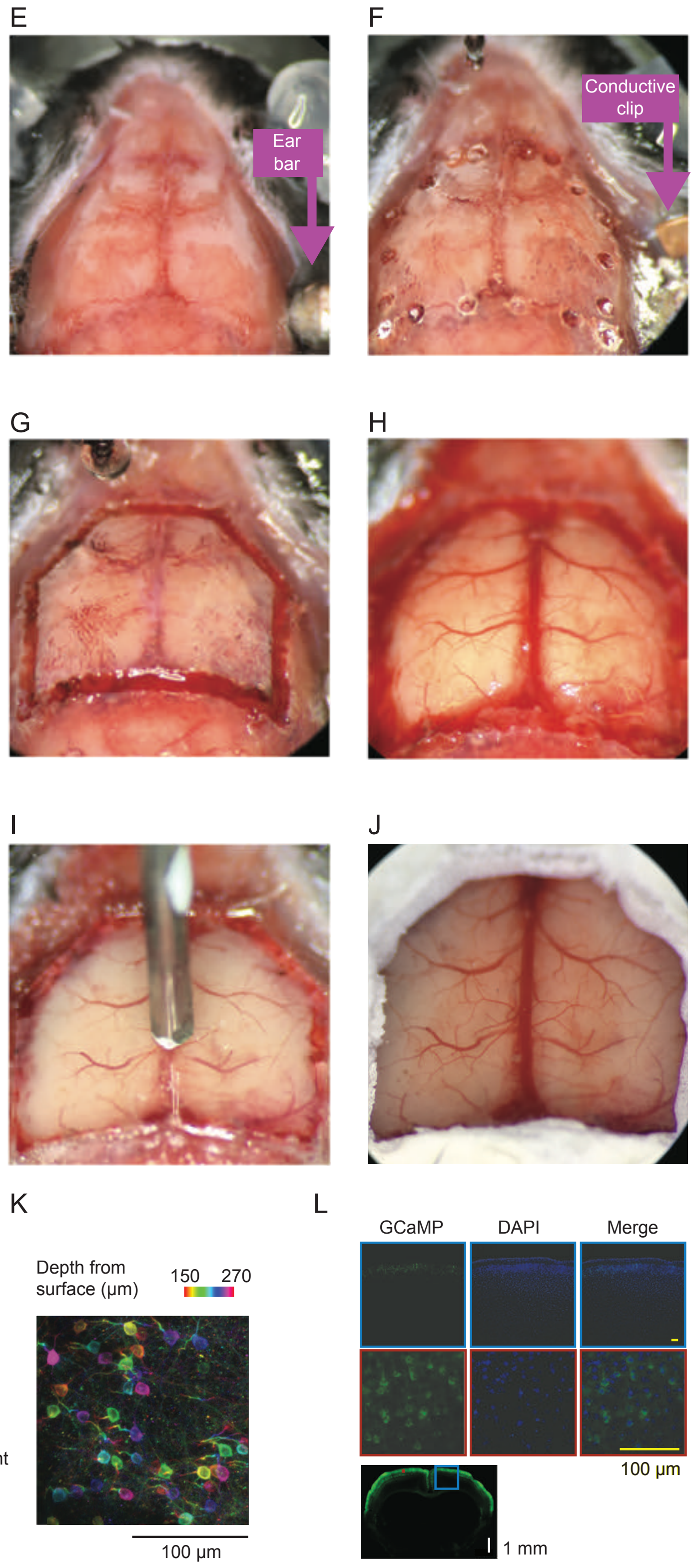
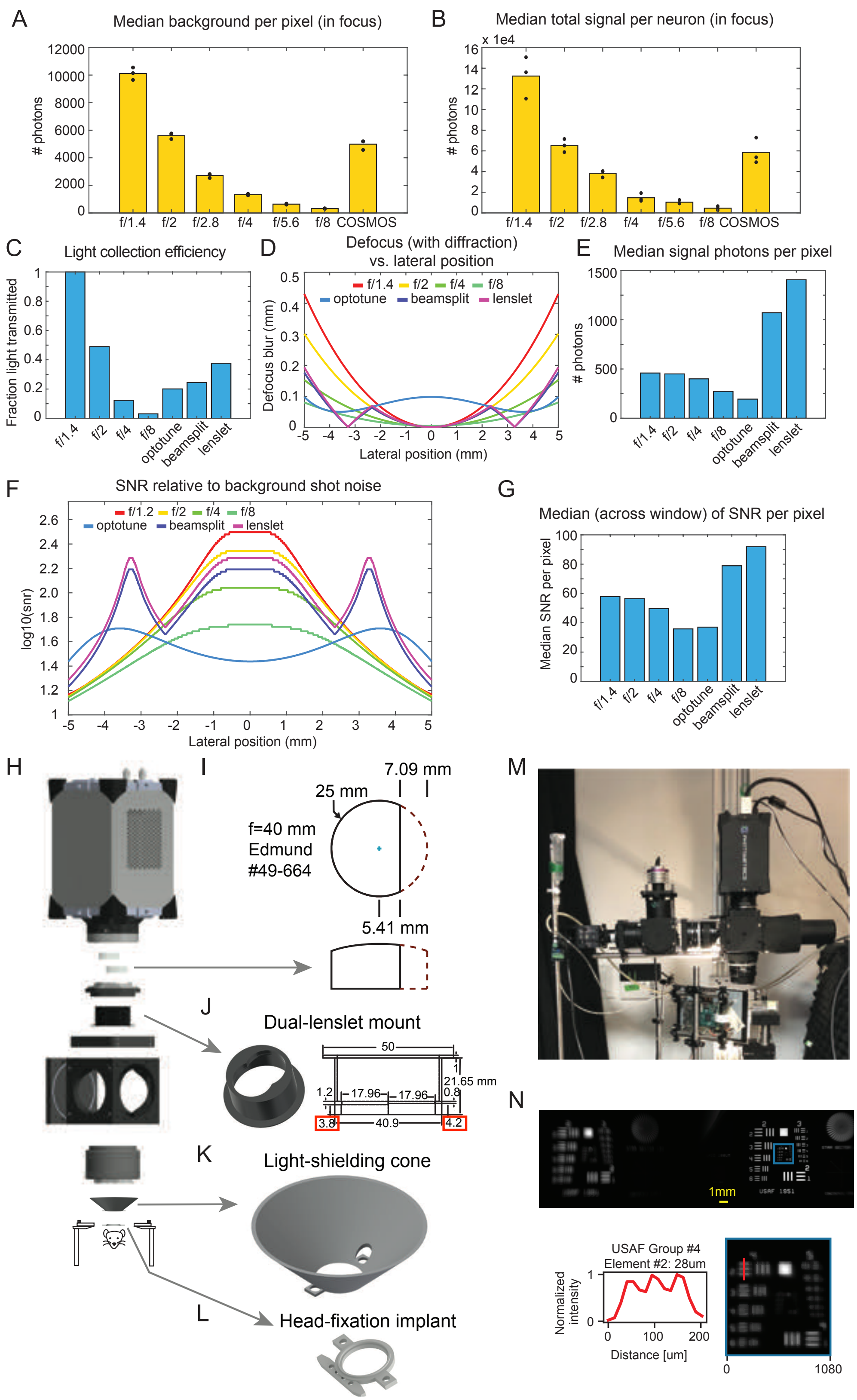


Figure S1 *(preceding page)*. Surgical procedure for implantation of large-scale window (relates to Figure 1).
(A) Geometry of curved glass window implant. (B) Screenshot of software interface for selecting keypoints. (C) Coordinates of keypoints in Panel B, with approximate thickness of the skull at each keypoint. Thickness varies for mice of different ages and sizes. (D) Photograph of robotic stereotax and setup. (E) After cutting skin back, ear bars are placed above skin and muscle, and skull is cleaned. (F) Conductive clip is attached to inside of the skin, and keypoints are drilled using robotic stereotax. Depth is determined using a conductivity-based autostop mechanism, or alternatively, inspection and knowledge of the rough depth of each keypoint based on previous mice. (G) Based on keypoints, robotic stereotax interpolates and automatically cuts full craniotomy. (H) After skull flap is loose, it is pulled off quickly using forceps and blood is cleaned using absorptive sponge swabs. (I) Curved glass implant is positioned using a blunt needle connected to vacuum, and pushed down such that glass is in contact with brain across the full window. (J) Window is cemented into position. (K) A depth-colored 120μm thick two-photon stack demonstrating density and health of the neurons. (L) High magnification histological section demonstrating GCaMP expression, with DAPI nuclear staining.



A
Median background per pixel (in focus)
photons
f/1.4 f/2 f/2.8 f/4 f/5.6 f/8 COSMOS
B
Median total signal per neuron (in focus)
x 1e4
photons
f/1.4 f/2 f/2.8 f/4 f/5.6 f/8 COSMOS
C
Light collection efficiency
Fraction light transmitted
f/1.4 f/2 f/4 f/8 optotune beamsplit lenslet
D
Defocus (with diffraction) vs. lateral position
f/1.4 f/2 f/4 f/8 optotune beamsplit lenslet
Defocus blur (mm)
Lateral position (mm)
E
Median signal photons per pixel
photons
f/1.4 f/2 f/4 f/8 optotune beamsplit lenslet
F
SNR relative to background shot noise
f/1.2 f/2 f/4 f/8 optotune beamsplit lenslet
log10(snr)
Lateral position (mm)
G
Median (across window) of SNR per pixel
Median SNR per pixel
f/1.4 f/2 f/4 f/8 optotune beamsplit lenslet
H
I
7.09 mm
25 mm
f=40 mm
Edmund
#49-664
5.41 mm
J
Dual-lenslet mount
50
21.65 mm
1.2
17.96
17.96
0.8
3.8
40.9
4.2
K
Light-shielding cone
L
Head-fixation implant
M
N
1mm
USAF Group #4
Element #2: 28um
Normalized intensity
Distance [um]
0
1080

Figure S2 ***(preceding page).*** **Theory, simulation, and design underlying the enhanced performance of COSMOS (relates to Figure 1).**
(A) Median background per pixel (across recovered neurons) for three recording epochs per detection design, from an in-focus ROI. Background is estimated using the CNMF-E algorithm. (B) Total signal per neuron (across recovered neurons) for three recording epochs per detection design, from an in-focus ROI. Signal is derived based on the spatial footprint and raw temporal traces estimated using the CNMF-E algorithm. (C) Simulated comparison of light collection efficiency of a conventional macroscope with different aperture settings, a tunable lens with a 16mm aperture, a multifocal two-camera beamsplitter design with f/2 objective aperture setting, and the multifocal dual lenslet design of COSMOS. Although the beamsplitter has a depth of field of f/2 it only has half the light collection. (D) Simulated defocus blur vs. lateral distance along the curved glass window (with 11 mm radius of curvature) for each design. (E) Simulated median (taken laterally across the window) of photon density for each design. (F) Simulated signal to noise ratio vs. lateral position for each design. Assumes a Gaussian point spread function. Secondary peaks are visible for the two multifocal designs. Although f/1.4 has a higher peak SNR when in focus, SNR drops off quickly as the image defocuses. (G) Simulated median (taken laterally across the window) of the SNR for each pixel. Lenslet outperforms the other designs. (H) CAD model of dual-lenslet system. (I) Specifications of an individual lenslet. 7.09 mm from the edge of an off-the-shelf aspherical achromatic lenslet was milled away. (J) Schematic and dimensions of dual-lenslet mount. (K) CAD model of light-blocking cone used to mitigate imaging artifacts induced by visual stimuli. (L) CAD model of headbar used for head-fixing the mouse. (M) Photograph of COSMOS microscope and behavioral setup. (N) Raw sensor image of USAF 1950 Resolution test chart.

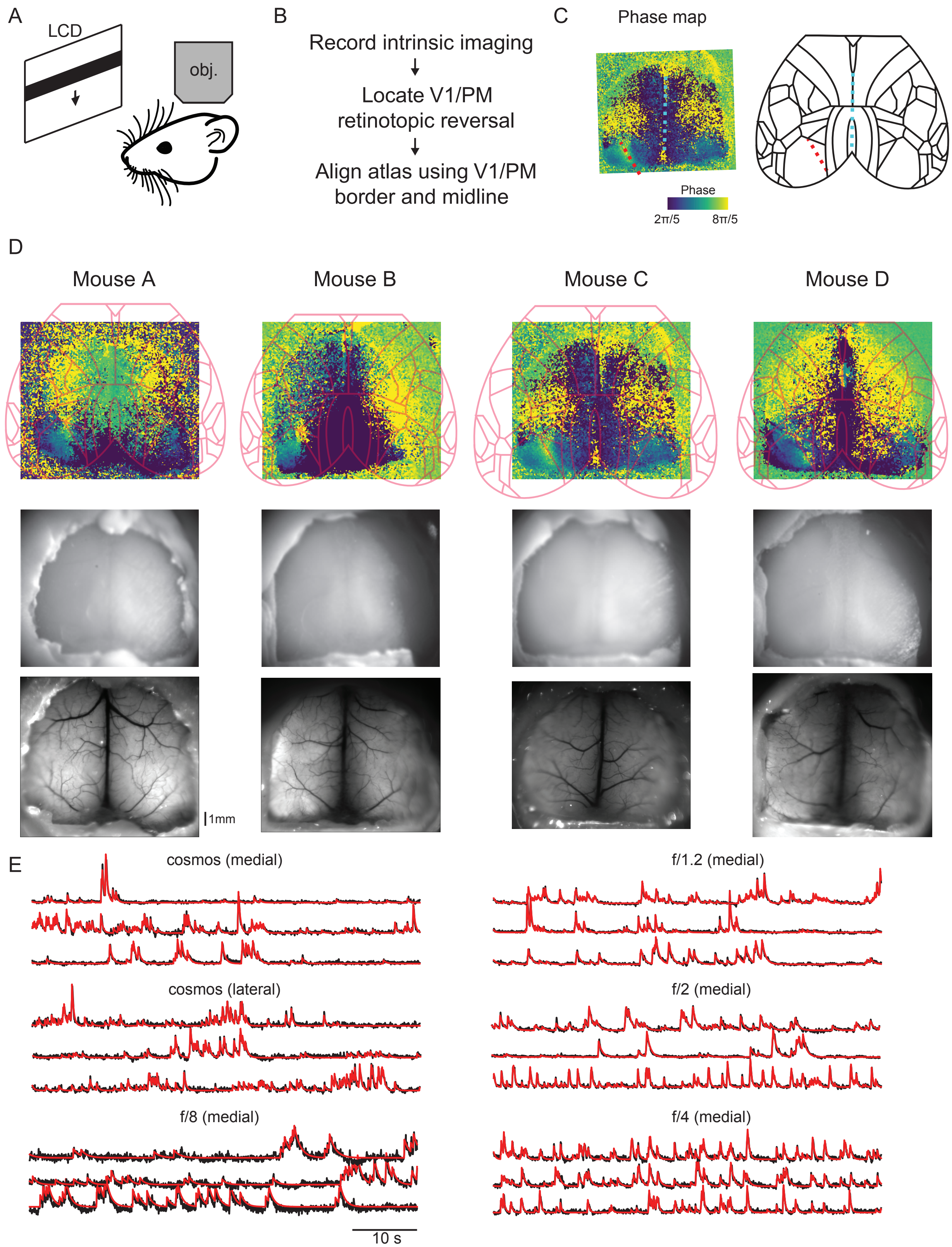

A
LCD
obj.
B
Record intrinsic imaging
Locate V1/PM retinotopic reversal
Align atlas using V1/PM border and midline
C
Phase map
Phase
2π/5
8π/5
D
Mouse A
Mouse B
Mouse C
Mouse D
1mm
E
cosmos (medial)
cosmos (lateral)
f/8 (medial)
f/1.2 (medial)
f/2 (medial)
f/4 (medial)
10 s

Figure S3 ***(preceding page)*****. Intrinsic imaging and atlas alignment method (relates to Figure 1).**
(A) Experimental setup for intrinsic imaging with a moving bar visual stimulus. (B) Flowchart of intrinsic imaging-based alignment procedure. (C) Example raw phase map with midline and V1/PM border overlaid, also indicated on the atlas. (D) (top) Alignment maps for all mice. (middle) The 700 nm illuminated images used for intrinsic imaging which were aligned to (bottom) the COSMOS fluorescence images. (E) Example raw traces with overlaid smoothed traces (both as output from CNMF-E algorithm) of a few example sources for different imaging conditions. The f/8 macroscope setting shows a clear degradation of signal-to-noise ratio in the raw traces.

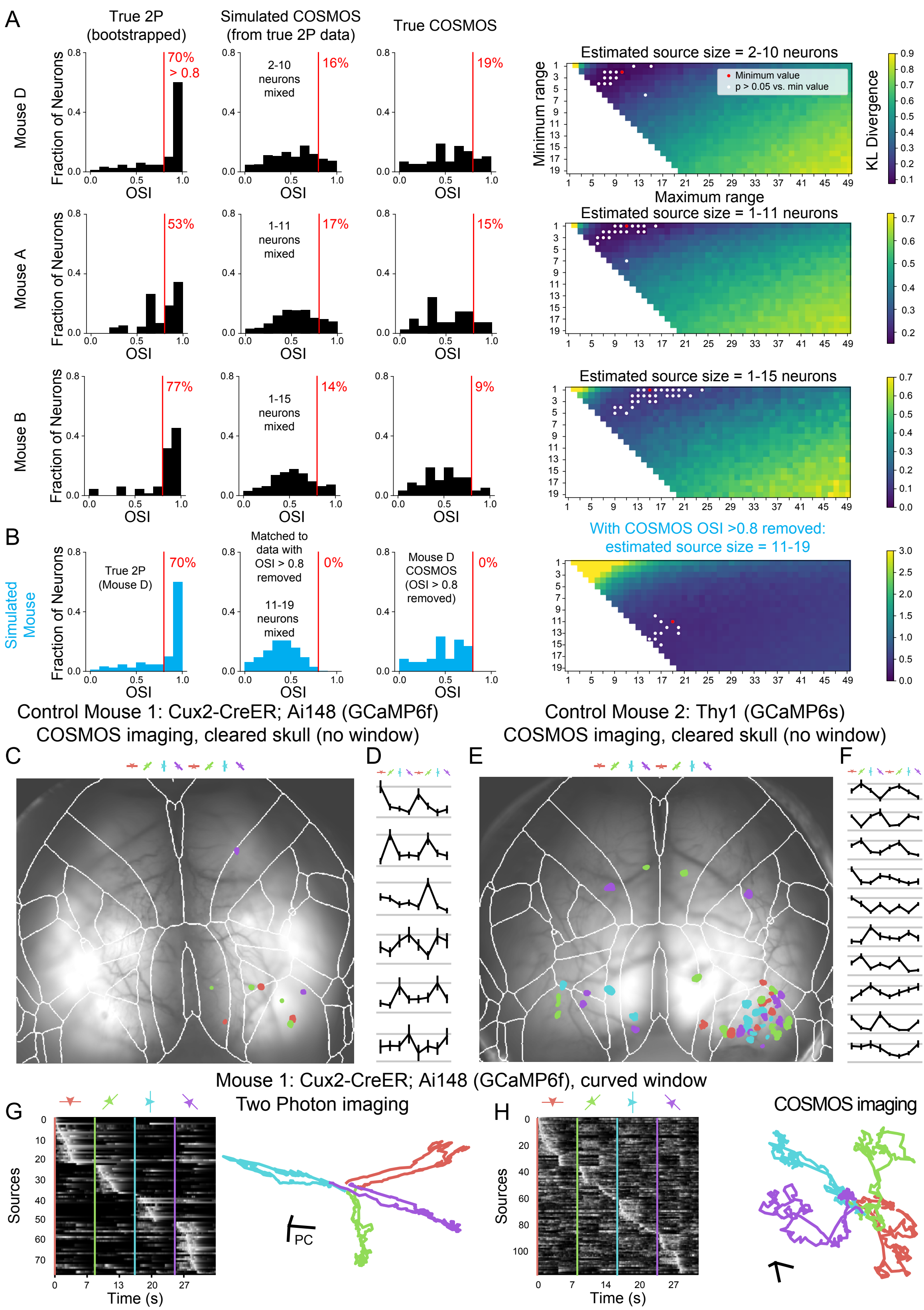

A
True 2P (bootstrapped)
Simulated COSMOS (from true 2P data)
True COSMOS
Mouse D
Mouse A
Mouse B
Fraction of Neurons
OSI
70% > 0.8
2-10 neurons mixed
16%
19%
53%
1-11 neurons mixed
17%
15%
77%
1-15 neurons mixed
14%
9%
Estimated source size = 2-10 neurons
Minimum value
p > 0.05 vs. min value
Minimum range
Maximum range
KL Divergence
Estimated source size = 1-11 neurons
Estimated source size = 1-15 neurons
B
Simulated Mouse
True 2P (Mouse D)
70%
Matched to data with OSI > 0.8 removed
11-19 neurons mixed
0%
Mouse D COSMOS (OSI > 0.8 removed)
0%
With COSMOS OSI >0.8 removed: estimated source size = 11-19
Control Mouse 1: Cux2-CreER; Ai148 (GCaMP6f)
COSMOS imaging, cleared skull (no window)
Control Mouse 2: Thy1 (GCaMP6s)
COSMOS imaging, cleared skull (no window)
C
D
E
F
Mouse 1: Cux2-CreER; Ai148 (GCaMP6f), curved window
Two Photon imaging
COSMOS imaging
G
H
Sources
Time (s)
PC

Figure S4 ***(preceding page).*** **Additional analysis of visual-stimulus source characterization (relates to Figure 2).**
(A) Simulation-based estimate of number of neurons per source. First column: Histograms of orientation selectivity index (OSI) for two-photon data shown in Figure 2F (bottom row) generated by resampling with replacement 500 times. Second column: on each of 500 iterations, groups of neurons (quantities chosen by sampling uniformly from the range indicated on the subplot) were averaged together. These 500 mixed sources then used to generate OSI histograms. Third column: true OSI histograms derived from COSMOS data. Red percentages denote the fraction of neurons with OSI > 0.8. Fourth column: similarity between simulated COSMOS OSI distributions and the true COSMOS distribution measured by KL Divergence (lower is more similar). The most similar range of sources to mix together is denoted by a red dot. White dots denote ranges that are statistically indistinguishable from the red dot range (one-sample T-test vs. model marked with red dot, $p > 0.05$). 10 simulated histograms were generated for each point in parameter space. Results shown for three mice. (B) Format matches the above panels using data from Mouse D but where the target COSMOS distribution has all sources with OSI > 0.8 removed. The best fitting simulated histogram corresponds to one with a larger degree of neuron mixing. (C) Atlas aligned to cleared skull of a Cux2-CreER; Ai148 (GCaMP6f) mouse. Highlighted sources were identified as visually responsive (one-way ANOVA, $p < 0.01$) to a drifting grating stimulus presented in eight directions. The orientation selectivity of each visually-responsive source is indicated in color. (D) Stimulus-averaged fluorescence responses to each grating direction from sources extracted from right visual cortex of the mouse from Panel C under the COSMOS microscope. Each source is normalized to its peak response. Error bars represent S.E.M. (E-F) Format matches Panels C-D but for a Thy1 (GCaMP6s) mouse imaged through its cleared skull. (G) Trial-averaged, max-normalized, fluorescence traces taken from visually responsive neurons found in right visual cortex of Mouse 1 (Cux2-CreER; Ai148, curved window) under a two-photon microscope. The right panel shows neural population responses for trial-averaged data shown in the left panel projected onto the basis vectors defined by PCA. (H) The same visual stimuli were presented to Mouse 1, under the COSMOS microscope, to reveal the trial-averaged, max-normalized fluorescence traces of sources taken from both visual cortical areas of Mouse 1 (left) and corresponding trajectories in PC space (right).

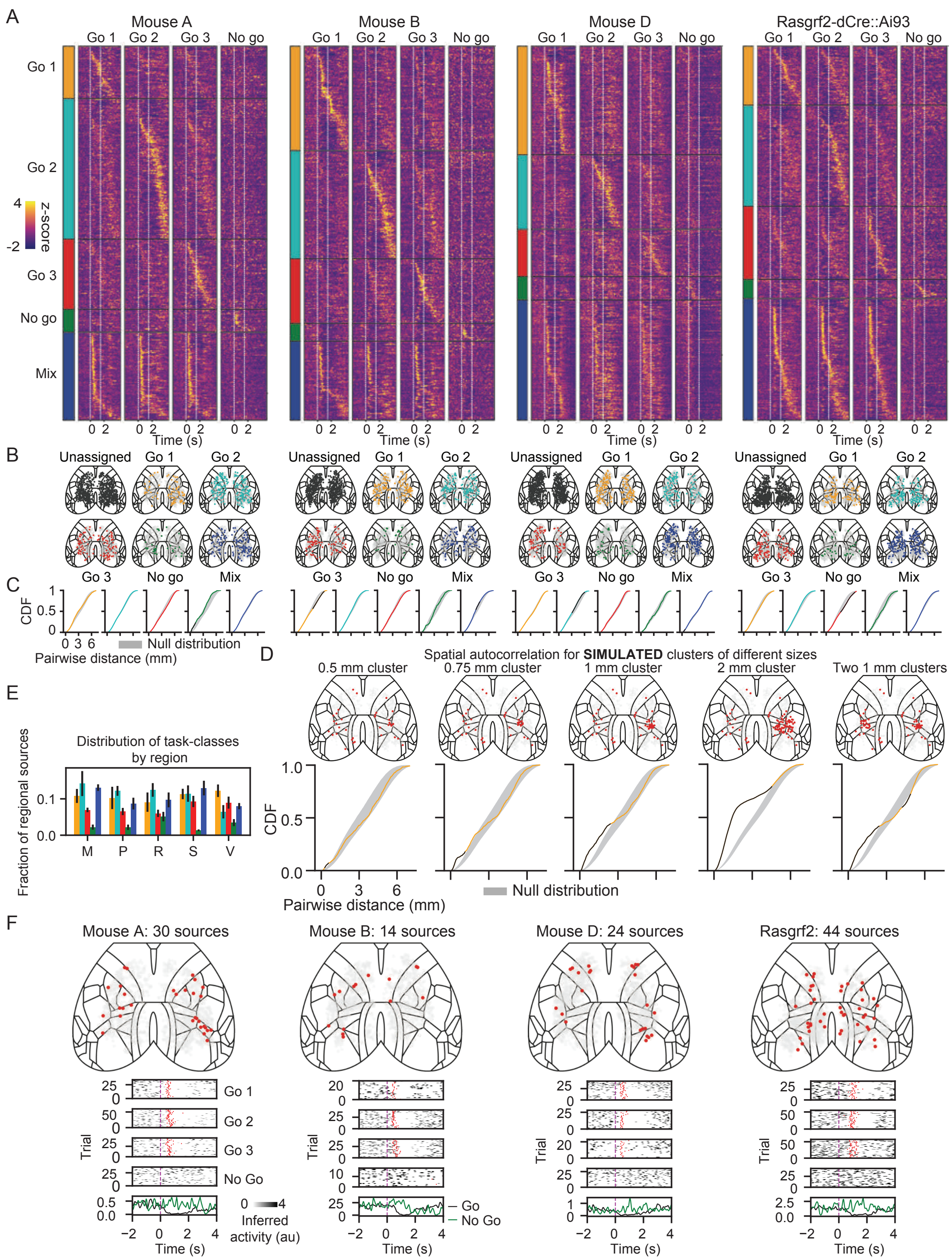

A
Mouse A
Mouse B
Mouse D
Rasgrf2-dCre::Ai93
Go 1
Go 2
Go 3
No go
Mix
z-score
4
-2
Time (s)
B
Unassigned
C
CDF
Pairwise distance (mm)
Null distribution
D
Spatial autocorrelation for **SIMULATED** clusters of different sizes
0.5 mm cluster
0.75 mm cluster
1 mm cluster
2 mm cluster
Two 1 mm clusters
E
Distribution of task-classes by region
Fraction of regional sources
M
P
R
S
V
F
Mouse A: 30 sources
Mouse B: 14 sources
Mouse D: 24 sources
Rasgrf2: 44 sources
Trial
No Go
Inferred activity (au)
Go
Time (s)

Figure S5 ***(preceding page).*** **Further details about task-responsive sources and their spatial distributions (relates to Figure 3).**

(A) Four additional mice, same format as Figure 3J, showing average traces of sources assigned to each of five task-related classes. This includes one Rasgrf2-2A-dCre:ai93, a different genetic line that labels a potentially non-identical subset of superficial neurons as Cux2-CreER. We see similar response profiles in both genotypes. We did not do any further analyses with the Rasgrf mouse data. (B) Spatial distribution of sources assigned to each task-related class. (C) Spatial autocorrelation analysis of distributions shown in Panel B (CDF is cumulative fraction of source-pairs separated by a given distance). Only five task classes had any significant deviations from the null distribution across all mice and distances (corrected $p < 0.05$, permutation test, significant values are indicated with black line overlay), and the significant task classes or CDF values were not consistent across mice. (D) Spatial autocorrelation analysis on simulated data to investigate sensitivity to clusters of different sizes. Clusters at least 1mm in diameter (1/3 of sources within the cluster area were assigned to the cluster) displayed clear significant ($p < 0.05$) differences from the null distribution (marked with black line overlay). (E) Distribution of task classes within each region (M=motor, P=parietal, R=retrosplenial, S=somatosensory, V=visual). Error bars are S.E.M. across n=4 Cux2-CreER mice. (F) Additional spatial distributions and examples of sources exhibiting “lick OFF” dynamics.

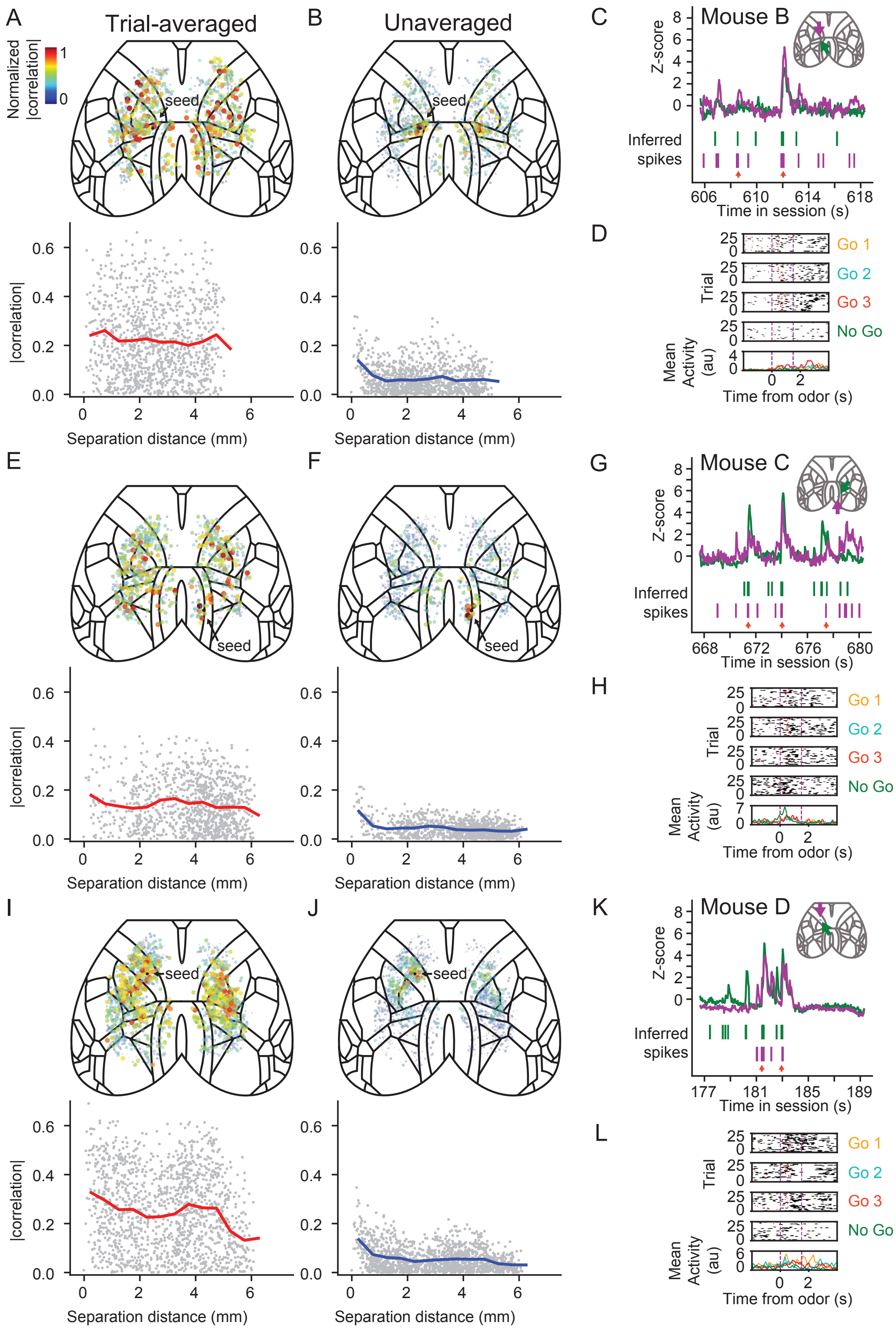


Figure S6. Additional data related to unaveraged correlation structure (relates to Figure 4).
(A) Trial-averaged activity correlations: (top) spatial distribution, (bottom) correlation vs. distance to the seed. Seed location is represented as a black dot in right visual cortex, referenced by the labelled arrow. (B) Unaveraged activity correlations: (top) spatial distribution), (bottom) correlation vs. distance to the seed. This source exhibits symmetrically bilateral localized correlation. (C) Example of unaveraged correlated activity between the seed source and a neighboring source (locations indicated on atlas inset). Z-scored fluorescence and inferred spikes are shown. Red arrows indicate timepoints when the seed source and its neighbor are active simultaneously. (D) Single trial responses of seed source. (E-H) Same, but for an additional mouse, with seed source in retrosplenial cortex. (I-L) Same, but for an additional mouse, with seed source in motor cortex.

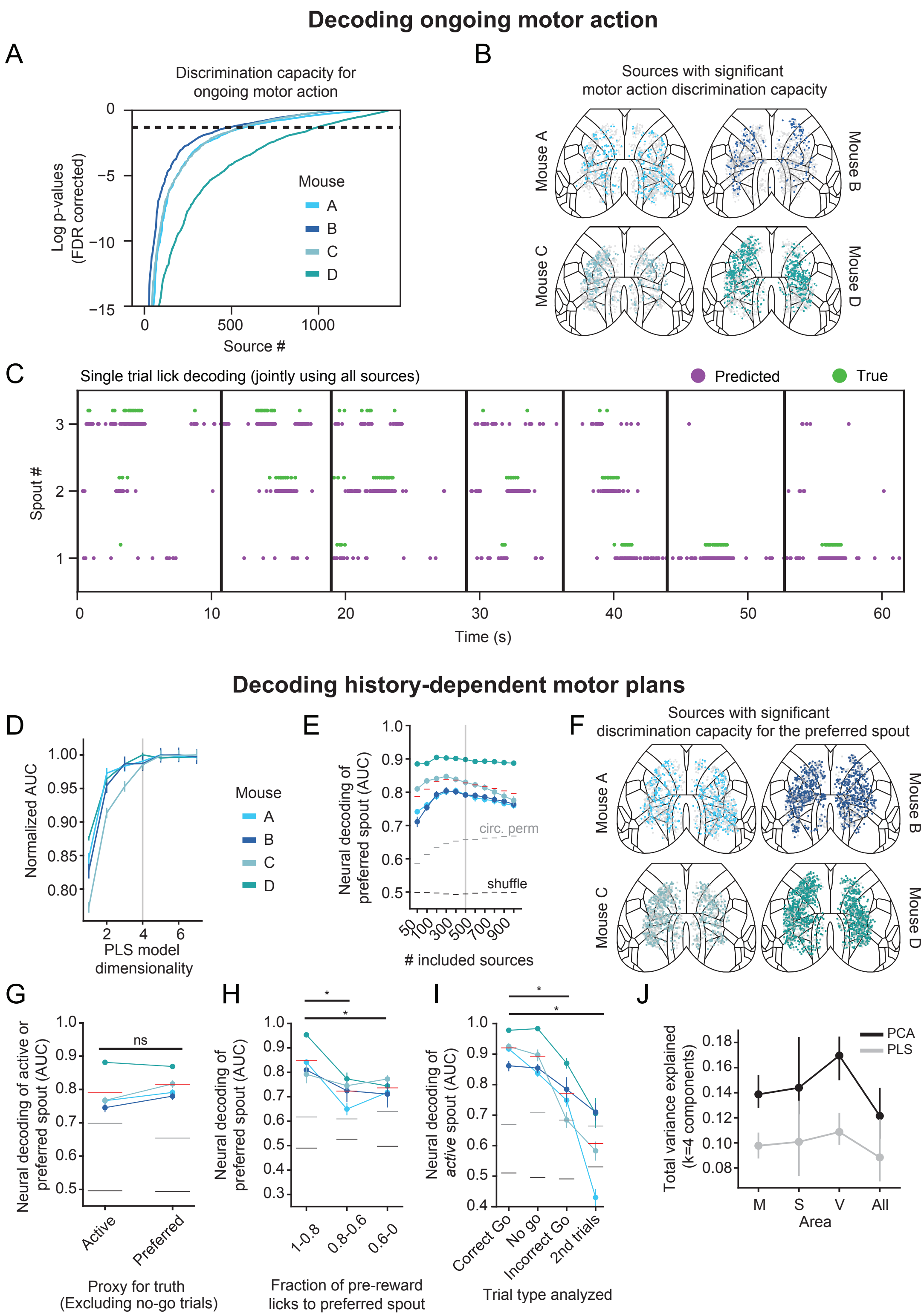

Decoding ongoing motor action
A
Discrimination capacity for ongoing motor action
Log p-values (FDR corrected)
0
−5
−10
−15
0
500
1000
Source #
Mouse
A
B
C
D
B
Sources with significant motor action discrimination capacity
Mouse A
Mouse B
Mouse C
Mouse D
C
Single trial lick decoding (jointly using all sources)
Predicted
True
Spout #
3
2
1
0
10
20
30
40
50
60
Time (s)
Decoding history-dependent motor plans
D
Normalized AUC
1.00
0.95
0.90
0.85
0.80
2
4
6
PLS model dimensionality
Mouse
A
B
C
D
E
Neural decoding of preferred spout (AUC)
1.0
0.9
0.8
0.7
0.6
0.5
circ. perm
shuffle
50
100
300
500
700
900
included sources
F
Sources with significant discrimination capacity for the preferred spout
Mouse A
Mouse B
Mouse C
Mouse D
G
Neural decoding of active or preferred spout (AUC)
1.0
0.9
0.8
0.7
0.6
0.5
ns
Active
Preferred
Proxy for truth (Excluding no-go trials)
H
Neural decoding of preferred spout (AUC)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
*
*
1-0.8
0.8-0.6
0.6-0
Fraction of pre-reward licks to preferred spout
I
Neural decoding of active spout (AUC)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
*
*
Correct Go
No go
Incorrect Go
2nd trials
Trial type analyzed
J
Total variance explained (k=4 components)
0.18
0.16
0.14
0.12
0.10
0.08
PCA
PLS
M
S
V
All
Area

Figure S7 ***(preceding page)*****. Additional analyses related to decoding motor actions and plans (relates to Figures 5-7).** (A) Ability of each neural source to discriminate between any of four behavioral states (licks towards one of the three spouts, or no lick). P-value computed for each source using Kruskal-Wallis H-test and FDR-corrected. Dashed horizontal line represents $p = 0.05$. (B) Spatial distribution of sources with significant motor action discrimination capacity ($p < 0.05$). (C) Example decoding of lick direction from neural activity at 30 Hz (on test dataset). Vertical lines represent trial onsets. (D) Model prediction accuracy versus the number of PLS basis vectors used. Five models were trained for each basis set size. (E) Pre-odor neural decoding performance as a function of number of neuronal sources used. Red lines indicate means across mice. Gray lines indicate performance of models trained and evaluated against circularly permuted true-spout labels. Black lines similarly represent models trained against randomly shuffled true-spout labels. Red lines indicate means across mice. Error bars show 99% bootstrapped confidence intervals. Vertical line at 500 indicates that all analyses that pooled across cortical regions used only the top 500 sources with most discrimination capacity for decoding to avoid model overfitting. (F) Spatial distribution of sources with significant discrimination capacity ($p < 0.05$) for the spout that was most-licked during the pre-reward period. P-values were computed for each source using a Kruskal-Wallis H-test and FDR-corrected. (G) Comparison of decoding performance of either the active spout (left) or the most-licked spout (right). Corrected paired t-test p-value shown between the active and preferred conditions. Note that trials where the active spout and most-licked spout were not identical were never used for the training of any PLS models. (H) Neural decoding of the preferred spout as a function of fraction of pre-reward licks towards the preferred spout. (I) Model AUC when evaluated on "correct go trials" (defined as those where at least 70% of total licks were to the active spout), "no go" trials, "2nd trials", and "incorrect go trials" ("go" trials where less than 70% total trial licks were towards the active spout). "Incorrect go trials" are defined differently here, versus the main text (where less than 30% was the criterion), because multiple unique trials towards each spout were needed to accurately compute the AUC. (J) The total variance explained, shown for the top four basis vectors computed using either PLS (bottom, gray) or Principal Components Analysis (PCA, top, black), averaged across all four mice.

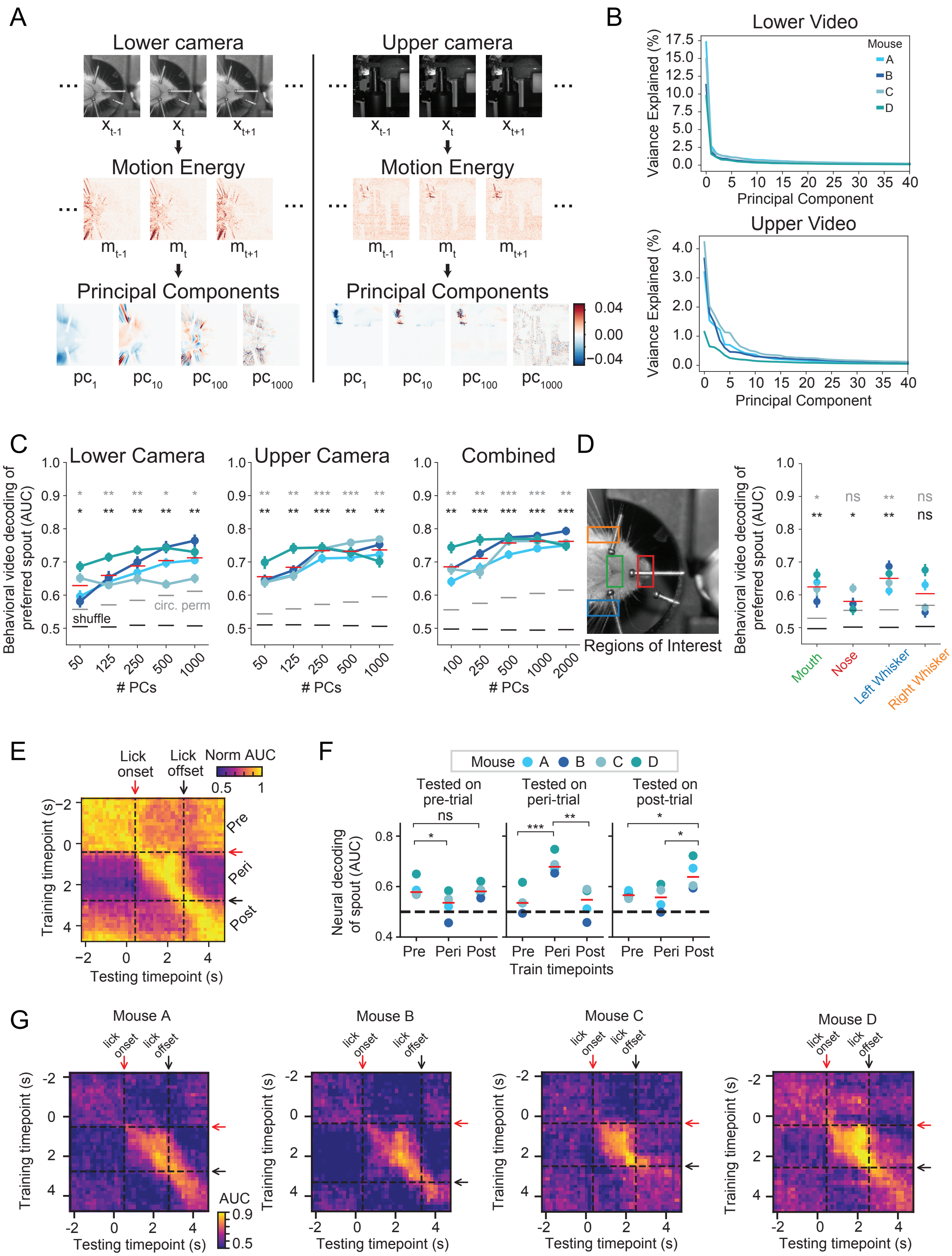

A
Lower camera
Upper camera
Motion Energy
Principal Components
0.04
0.00
−0.04
B
Lower Video
Upper Video
Mouse
A
B
C
D
Variance Explained (%)
Principal Component
C
Lower Camera
Upper Camera
Combined
Behavioral video decoding of preferred spout (AUC)
shuffle
circ. perm
PCs
D
Regions of Interest
Behavioral video decoding of preferred spout (AUC)
Mouth
Nose
Left Whisker
Right Whisker
E
Lick onset
Lick offset
Norm AUC
Training timepoint (s)
Testing timepoint (s)
Pre
Peri
Post
F
Mouse A B C D
Tested on pre-trial
Tested on peri-trial
Tested on post-trial
Neural decoding of spout (AUC)
Train timepoints
G
Mouse A
Mouse B
Mouse C
Mouse D
lick onset
lick offset
AUC
Training timepoint (s)
Testing timepoint (s)

Figure S8 *(preceding page)*. Active spout decoding from high speed videos of mouse behavior and cross-temporal neural decoding (relates to Figures 6 and 7).

(A) Extracting motion-related features from high-speed behavior videos. (B) Variance explained by behavioral video motion energy principal components. (C) Results from PLS decoding of active spout from pre-odor video motion features. (D) Results of decoding using pre-odor motion features from specific regions of interest. (E) We asked if there might be a difference between widespread neural representations of motor plans, and of motor plan execution. To do this, we compared sets of PLS models trained on different epochs of the trial. By training models on a single time bin (e.g. during the pre-odor period) and evaluating prediction performance on a different time bin (e.g. during movement), we tested the similarity of the PLS models over the trial. The goal was to understand whether a single projection of the data might best discriminate the trial types, or whether cortex might represent motor plans during the intertrial interval differently from actual directed licking during the peri-movement period (between lick onset and lick offset). Examining how well each of the time-specific models predicted other time bins revealed clear structure. In grouping these time bins into pre-, peri-, and post-task epochs based on the average time of lick onset and offset, we found a dissociation between models trained on intertrial interval data (from the pre- and post-task epochs) versus those trained using peri-task period data. Specifically in this panel, AUC was computed for PLS bases trained to predict the active spout position from neural data in one time window, and then tested on neural data from all time windows and each mouse. AUC is normalized by the maximum AUC for each testing time window, and averaged across four mice. Dashed lines represent the average time of lick onset and offset within the trial. For this analysis (panels E-G), testing was only done on trials where the active spout and preferred spout were identical (trials where the active spout and preferred spout were not the same were always excluded from training in all analyses). (F) Quantification of results in Panel E for individual mice, comparing cross-temporal prediction performance of three distinct time blocks: pre-task, peri-task, and post-task. For each testing time block, we statistically compared the mean prediction performance of bases trained on each time block (corrected t-test). These results are consistent with the interpretation that motor plans and consequent movements are represented by two distinct, but related, neural population representations. (G) Unnormalized cross-temporal AUC for each mouse. ns denotes corrected $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

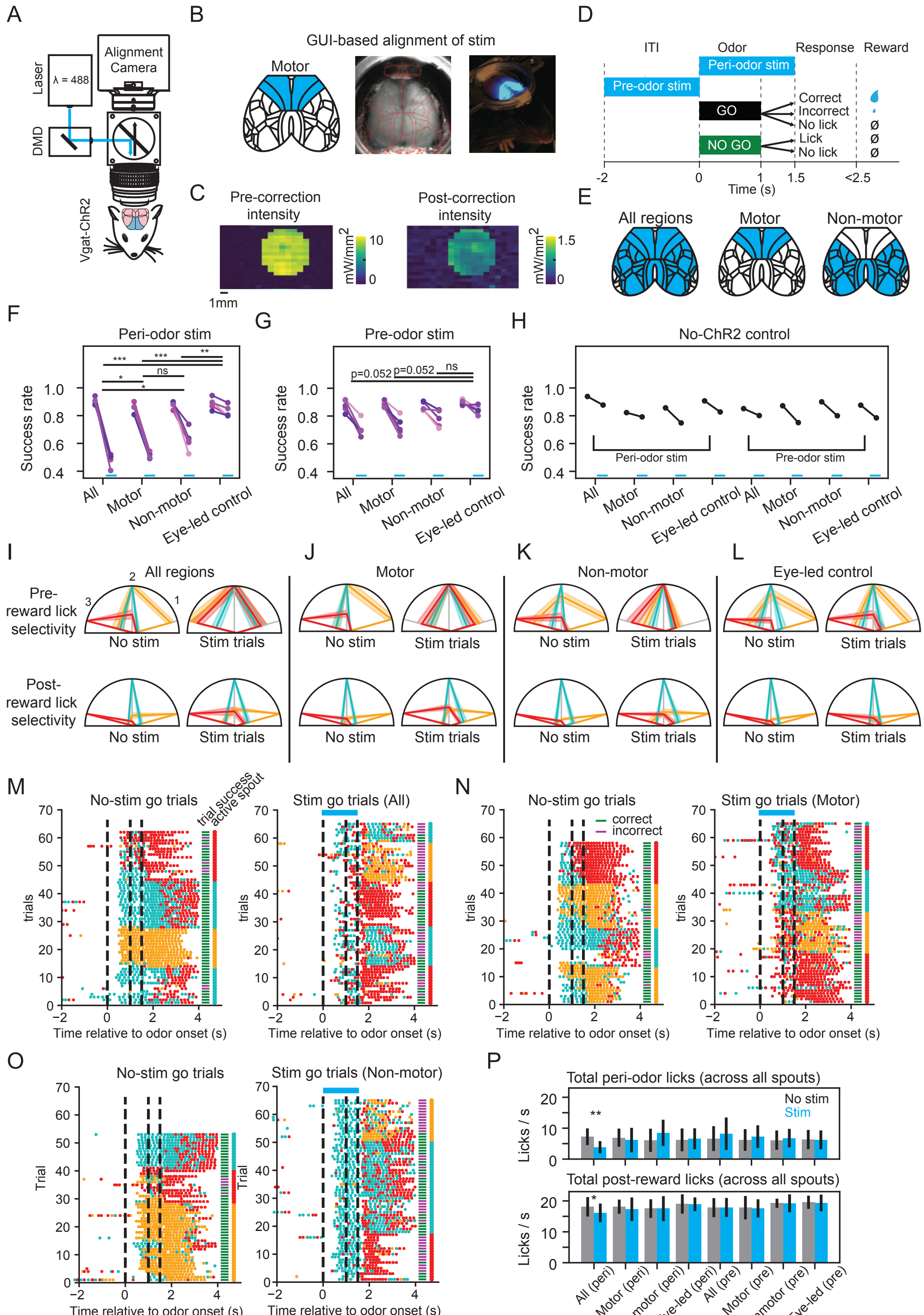

A
Laser
λ = 488
Alignment Camera
DMD
Vgat-ChR2
B
GUI-based alignment of stim
Motor
C
Pre-correction intensity
Post-correction intensity
mW/mm2
1mm
D
ITI
Odor
Response
Reward
Peri-odor stim
Pre-odor stim
GO
NO GO
Correct
Incorrect
No lick
Lick
No lick
Time (s)
E
All regions
Motor
Non-motor
F
Peri-odor stim
Success rate
All
Motor
Non-motor
Eye-led control
G
Pre-odor stim
p=0.052
ns
H
No-ChR2 control
Peri-odor stim
Pre-odor stim
I
All regions
Pre-reward lick selectivity
Post-reward lick selectivity
No stim
Stim trials
J
Motor
K
Non-motor
L
Eye-led control
M
No-stim go trials
trial success
active spout
Stim go trials (All)
trials
Time relative to odor onset (s)
N
correct
incorrect
Stim go trials (Motor)
O
Stim go trials (Non-motor)
Trial
P
Total peri-odor licks (across all spouts)
No stim
Stim
Licks / s
Total post-reward licks (across all spouts)
All (peri)
Motor (peri)
Non-motor (peri)
Eye-led (peri)
All (pre)
Motor (pre)
Nonmotor (pre)
Eye-led (pre)

Figure S9 ***(preceding page).*** **Inhibition of non-motor areas impairs lick-to-target task performance (relates to Figures 5-7).**
(A) In addition to recording neuronal activity, we tested if the task-activity we found distributed across many non-motor areas was causally involved in task performance. We investigated this because it is known that specific areas of cortex can modulate the firing of neurons in primary and secondary motor cortex. For example, projections from S1 can drive M1 activity and drive the initiation of whisking (Sreenivasan et al., 2016). But in addition, posterior regions like retrosplenial cortex and primary sensory areas (Barthas and Kwan, 2017; Yamawaki, Radulovic, and Shepherd, 2016; Zingg et al., 2014), also project to secondary motor cortex and may thus play an important role in producing any dynamics observed in motor areas. We therefore set out to study whether mice could accomplish our task without normal input from all of these non-motor areas by simultaneously inhibiting multiple cortical regions using the Digital Micromirror Device (DMD) system shown here. This approach avoided the possibility that uninhibited areas in cortex might compensate for the acute shutdown of a single area. (B) Atlas alignment. (C) Intensity calibration. (D) Stimulation protocol. (E) Stimulation patterns. (F) Success rate on peri-odor stimulation vs. no stimulation trials, for each stimulation pattern, n=5 mice, corrected paired t-test between stimulation patterns. (G) Success rate on pre-odor stimulation vs. no stimulation trials, for each stimulation pattern, n=5 mice, corrected paired t-test between stimulation patterns. (H) Success rate on peri- and pre-odor stimulation for a mouse not expressing ChR2. (I-L) Summary of lick selectivity averaged across n=5 mice (with S.E.M. shown) after odor delivery, but before (top) or after (bottom) reward delivery during non-stimulation trials or stimulation trials, for each stimulation pattern. Colors denote the normalized number of licks towards each spout on trials when a given spout is active. (M-O) Example lick rasters demonstrating that licking is not fully abolished during photoinhibition, but rather the specificity of it is disrupted (for each stimulation pattern, some sessions showed incorrect preference for a single spout, while others showed nonspecific licking towards multiple spouts). (P) Comparison of total number of peri-odor and post-reward licks during no-stimulation (gray) and stimulation (blue) trials (mean +/- S.E.M. across mice, corrected paired t-test between stimulation patterns). ns denotes corrected $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.