Concerted modulation of spontaneous behavior and time-integrated whole-brain neuronal activity by serotonin receptors

Drew Friedmann[1,2,8], Xavier Gonzalez[3], Ashley Moses[1], Tanner Watts[1,3,4], Anthony Degleris[3,5], Nicole Ticea[3,6], Jun H. Song[1], Sandeep Robert Datta[7], Scott W. Linderman[3,8]*, Liqun Luo[1]*

[1]Department of Biology, Howard Hughes Medical Institute, Stanford University
[2]Cajal Neuroscience Inc.
[3]Department of Statistics, Wu Tsai Neurosciences Institute, Stanford University
[4]Department of Computer Science, University of Utah
[5]Department of Electrical Engineering, Stanford University
[6]Department of Physics, Stanford University
[7]Department of Neurobiology, Harvard Medical School
[8]Equal contribution
*Corresponding authors. Email: scott.linderman@stanford.edu (SWL), lluo@stanford.edu (LL)

ABSTRACT

Serotonin neurons from the raphe nuclei project across the entire brain and modulate diverse physiology and behavior by acting on 14 distinct receptors. Here, we take a step dissecting this complex process by examining the effects of agonists and antagonists of four widely expressed serotonin receptors (2A, 2C, 1A, and 1B) on spontaneous mouse behavior. We also analyzed time-integrated whole-brain neuronal activity of these same mice based on intensity of Fos, product of an immediate-early gene. Low-dimensional representations of behavioral and Fos map data shed light on the dominant factors of variation in each domain, captured predictable differences across drug groups, and enabled predictions of behavioral changes following perturbations in Fos maps and vice versa. Our study provides rich resources on the effects of manipulating serotonin receptors on animal behavior and whole-brain integrated neuronal activity. It also establishes an experimental and analysis paradigm for interrogating the relationship between behavior and neuronal activity across different time scales.

INTRODUCTION

A central goal of neuroscience is to link animal behavior with neuronal activity in the brain. This has traditionally been approached by focusing on a specific behavior and neuronal activity in a specific brain region. Much progress has been made by such focused studies, from sensory perception, decision making, and cognitive maps to motor planning and action[1–4]. With recent technical advances, it is now possible to record or optically image activity of thousands of neurons simultaneously in behaving animals[5,6]. Substantial advances have also been made in methods that record and quantify spontaneous behavior of animals unbiased to any training[7–10]. These advances create an opportunity for investigating the relationship between large-scale neuronal activity and natural behavior in the same animal.

Indeed, in simpler nervous systems with optical transparency, such as *C. elegans* and zebrafish larvae, recent efforts have made it possible to perform optical imaging of most neurons in behaving animals[11–14]. However, measuring whole-brain activity with single-neuron resolution in a behaving mammal is not yet possible and will unlikely be in the foreseeable future. Large-scale optical imaging can only access neurons close to the surface of the skull, and high-density electrophysiological probes can only access but a tiny fraction of the many millions of neurons in a mammalian brain. An alternative to real-time neuronal activity is to measure time-integrated activity via the expression of immediate early genes (IEGs), which is rapidly and transiently induced by neuronal activity[15]. IEGs have been used to discover neuronal populations related to many specific behaviors and to manipulate neuronal populations that were previously activated by specific stimuli or behavioral tasks[16–19]. However, it is unclear the extent to which natural behavior relates to time-integrated activity in the whole brain.

Serotonin is a powerful neuromodulator of physiology in health and disease and is therefore a natural candidate for interventions that interrogate whole brain activity and behavior. In the mammalian brain, serotonin neurons reside in the raphe nuclei in the brainstem but project their axons across the entire brain[20,21]. Diverse functions have been ascribed to neuromodulation by serotonin, from sensory gating and locomotor activity to learning, social interactions, and regulation of diverse brain states such as anxiety and mood[22,23]. The diverse effects of serotonin on target neurons are mediated by 14 distinct serotonin receptors that are expressed in specific brain regions and cell types[24–26]. Except for serotonin receptor 3, which is an ionotropic receptor, all serotonin receptors are metabotropic receptors that are coupled to specific G proteins, causing either depolarization (e.g., serotonin receptor 2 subfamily, coupled to G_q/G_{11}) or hyperpolarization (e.g., serotonin receptor 1 subfamily, coupled to G_i/G_o) of target neurons[25]. Because of their potential medical applications (e.g., treating neuropsychiatric disorders), agonists and antagonists for preferential targeting of individual serotonin receptors have been extensively characterized[25,26].

Here we used a panel of well-characterized serotonin receptor agonists and antagonists to investigate their effects on spontaneous behavior of mice and on whole-brain integrated activity of the same mice via the expression levels of Fos, an IEG product. We developed statistical models to probe the relationship between behavioral 'syllables' and Fos maps and established a platform of predicting changes in behavior following perturbations in Fos maps and vice versa. Our study provided rich resources on how spontaneous behavior and integrated whole-brain activity are modulated by activation and inactivation of specific serotonin receptors. It also establishes a paradigm with which to investigate the relationship between behavior and brain activity under specific conditions in an unbiased manner.

RESULTS

We administered serotonin agonists/antagonists to 190 mice (9 groups, 19 mice per group) or saline (1 group of 19 mice) 30 min prior to monitoring their spontaneous behavior in an open arena. After 60 min of behavioral recording and 45 min of rest, mice were sacrificed for whole-mount Fos immunostaining (**Fig. 1a**). We first present our data and analyses on the effect of drugs on behavior (**Fig. 1** and **Fig. 2**). We then present our data and analyses on the effect of drugs on whole-brain Fos levels (**Fig. 3–5**). We finally investigate the relationship between behavior and Fos maps (**Fig. 6**). The data described below came from 188 mice that passed the quality control for behavioral data (**Fig. 1** only) and 168 mice that passed the quality control of data for both behavior and Fos map (**Fig. 2–6**; **Methods**).

Of the 14 serotonin receptors encoded by the mouse genome, we selected agonists and antagonists for their preferential activity on serotonin receptors 1A, 1B, 2A, and 2C (**Table 1**) because of their broad expression in the brain[24–26]. In the following text, we abbreviate the agonists as 1A+, 1B+, 2A+, and 2C+, and antagonists as 1A–, 1B–, 2A–, and 2B–, respectively. A 9th drug, 2AC+, is a mixed agonist for both serotonin receptors 2A and 2C but has historically been a popular choice as a 2A agonist. Drug concentrations were chosen based on prior characterization of their pharmacological properties[27–34].

Effects of drugs on spontaneous behavior

To determine the effect of serotonin receptor agonists and antagonists on behavior, we placed the mice in a circular open field and recorded spontaneous behavior of freely moving mice for 60 min. During the period between 30 and 45 min, a collimated white LED light illuminating the center of the arena was turned on, intending to heighten the anxiety of the mice (**Fig. 1a**).

For control mice injected with saline, examination of their trajectories (**Fig. 1b**, **Extended Data Fig. 1a**) and speed (**Fig. 1c, d**) revealed that their highest locomotion activity occurred during the first 15 min (and particularly the first 5 min) in the circular arena, likely reflecting their exploration of this novel environment. Mice slowed down during the 2nd 15-min period, increased locomotion in response to LED-on stimulation during the 3rd 15-min period, and returned to a low-activity state in the final 15-min period. Most drug cohorts followed a similar trend. 1A+ treated mice (1A+ mice hereafter) were an exception. They exhibited reduced movement at the outset and did not significantly increase speed in response to LED-on (**Fig. 1c**, quantified in the bottom panels). By contrast, 1B+ mice exhibited increased locomotion throughout the observation period compared to control;

they also showed an increase of locomotion in response to LED stimulation (**Fig. 1b–d**). Observation and quantitative analyses of spatial location revealed that all mice preferred to reside or move in the periphery, with 1B+ mice showing the strongest preference and 2C+ mice showing the largest shift to the periphery after LED onset (**Fig. 1b, e, Extended Data Fig. 1**).

Unsupervised clustering of behavior using the motion sequencing (MoSeq) paradigm[9,10,35] of all 188 mice during the 60-min period produced 50 "syllables" (**Table S1, Extended Data Fig. 2**). Spectral clustering of the MoSeq syllables produced hierarchical clusters that matched with independent human annotations of these syllables as "rear", "pause", "groom", and "walk/run" (**Fig. 1f**). Comparing the frequency of syllable use across different drug conditions revealed an increased use of walk/run syllables and a decreased use of pause syllables for 1B+ mice (**Fig. 1g**), consistent with increased locomotion speed from our scalar analysis above (**Fig. 1c**). By contrast, 1A+ mice exhibited a substantial increase of most pause syllable usage and decrease of usage of most walk/run syllables (**Fig. 1g**). Interestingly, 2A– mice exhibited similar syllable usage in pause and walk/run as 1A+ mice (**Fig. 1g**), suggesting shared behavioral effect for inhibiting 2A receptor and activating 1A receptor. As will be described later, 2A– and 1A+ also caused similar neuronal activation patterns in specific brain regions.

We also analyzed how drugs affected syllable use over time. **Fig. 1h** presents usage distributions of all 50 syllables across the 60-min period for all mice. To simplify visualization of this complex dataset, we identified the most-frequently-used syllable cluster in rolling bins of 200 frames (~6.7 seconds) across mice of each drug condition and displayed these across the temporal epochs (**Fig. 1i**). This analysis revealed that LED onset caused an increased use of run syllables in saline and most drug conditions. However, for mice treated with the 1A+ and 2A–, rearing appeared to be the dominant behavioral mode after LED onset (**Fig. 1i**).

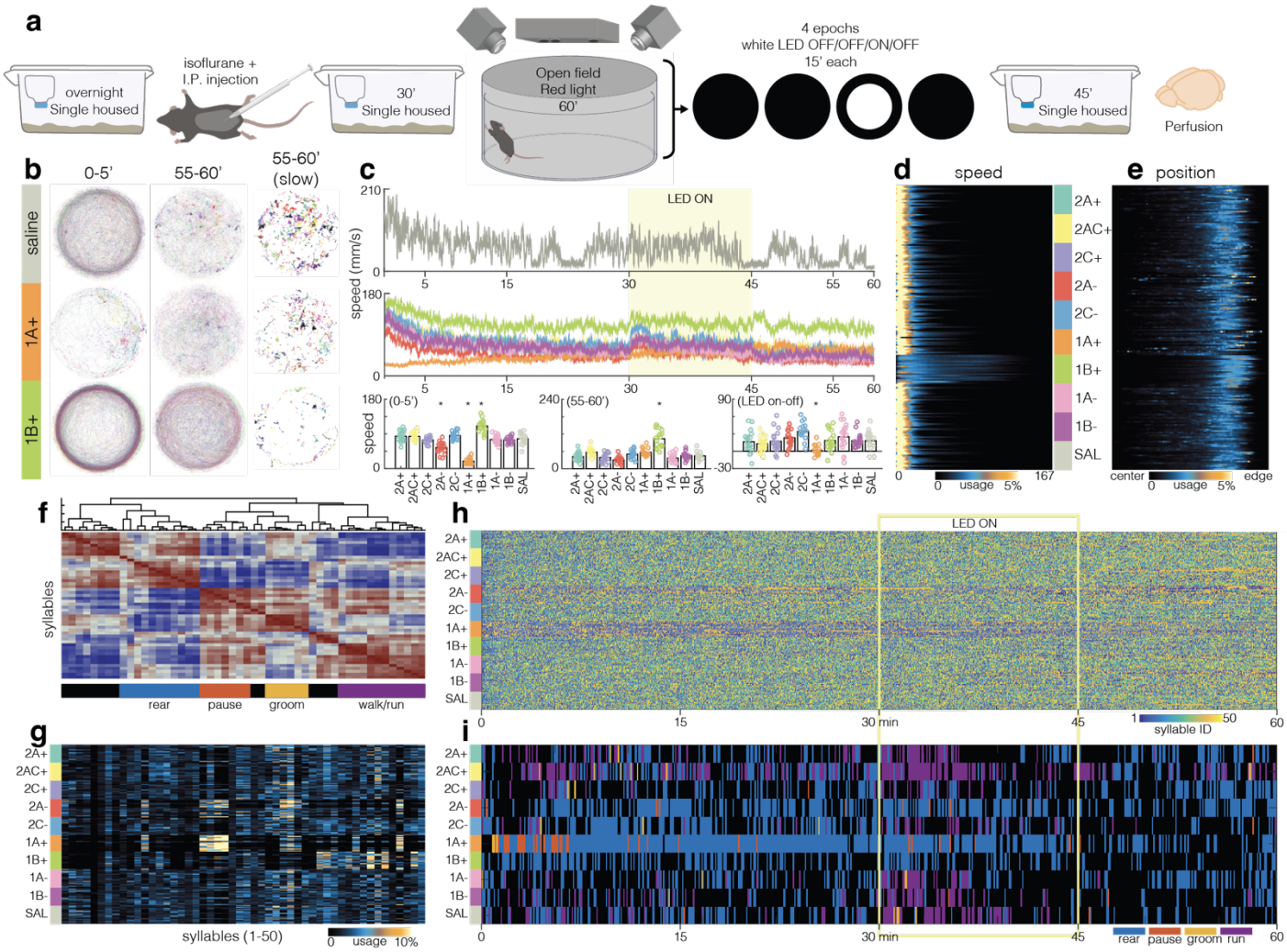


Fig.1 | The effect of serotonin receptor agonists and antagonists on spontaneous behavior. a, Experimental design and timeline for drug dosing, open field behavioral recording (at 30 frames/sec), and histology for whole-brain Fos mapping. **b,** Overlaid trajectories of individual mice from 3 of the 10 drug conditions. All mice from these conditions (n = 19, 17, 19) are shown for the displayed time windows. The 3rd column highlights trajectory positions with speed below 30 mm/s, highlighting that 1B+ reduces pauses. **c,** Top, smoothed 2D-speed recording for a single saline-treated mouse. The 15-min LED-on period is highlighted. Middle, average speed for each drug treatment group (color code as in **d**). The LED-on epoch is highlighted in yellow. Bottom (left and center), time binned average speed for each drug group during the first and last five minutes. Bottom right, the difference between the period 2 min after LED onset compared to 4 min prior. Group different from saline: * $p < 0.05$; ANOVA followed by Turkey post hoc test. **d,** Speed distribution for each mouse (rows), grouped by drug and measured as percentage of frame transitions representing each speed (mm/sec). **e,** Positional distribution for each mouse (rows), grouped by drug and measured as a percentage of frames spent at a given concentric position of the arena; 140 bins. **f,** Cladogram output of a spectral clustering of MoSeq syllables highlights groups of similar structure. Manual descriptions of blinded video from each syllable resulted in categorical descriptors for four major tree branches of the hierarchy. **g,** Average usage (as a percentage of frames) of each syllable for all mice across drug groups aligned to the cladogram in **f**. **h,** Colormap of MoSeq syllables 1–50 shown for all mice (n = 188) and frames (n = 108000). Conditions 1A+ and 2A– stand out as horizontal bands, but all groups shift in syllable usage across time (blue to yellow). **i,** Differential usage of the syllables comprising the four major categorical behaviors as described in **f** by drug group across the full recording session. Calculated as the modal category in rolling bins of 200 frames. LED response is seen as an increase in the walk/run syllables in the 2A+, 2AC+, 1A–, 1B–, and saline conditions.

Decomposition of behavior

Typical MoSeq analyses consider egocentric pose and do not take absolute position or time into account. To understand how syllable usage varies with position in the arena and the time elapsed since the start of the session, we used nonnegative Tucker decomposition[36,37] (**Methods**). First, we organized the data for each mouse into a 3D tensor where each entry specifies the number of video frames in which a particular syllable was performed at a specific location within a certain time epoch. The Tucker decomposition approximates each mouse's behavioral data as a weighted combination of shared factors. Like in a singular value decomposition of a matrix, each shared factor is a low-rank combination of *syllable*, *position*, and *epoch factors*, which specify what, where, and when syllables are used, respectively. Finally, the *mouse weights* specify how to combine these shared factors to best approximate each mouse's behavioral tensor. We used cross-validation to determine that 22-dimensional mouse weights, 22 syllable factors, 4 position factors, and 4 epoch factors best explain the data (**Fig. 2a, Extended Data Fig. 3, Fig. 4a–c**).

The epoch, position, and syllable factors capture low-dimensional structure in syllable usage (**Fig. 2b**). For example, the first epoch factor shows an exponential decay over the first 20 min of the session, which captures transient behavior after mice were initially placed in the arena (**Fig. 2b,** top left, red line). The second epoch factor ramps up as the mice become accustomed to the environment; the third captures the sudden change when the lights turn on; and the fourth captures the last 15 min when the lights are turned off again (**Fig. 2b,** top left). The position factors tile locations near the center, middle, and periphery of the arena, allowing the model to capture different syllable usage at different locations (**Fig. 2b,** bottom left). The syllable factors capture sets of syllables that tend to be used in similar locations and time epochs, like subsets of pausing syllables (**Fig. 2b,** right).

The mouse weights offer a condensed summary of each mouse's syllable usage across space and time (**Extended Data Fig. 4d**). The weights show considerable variability across drug groups, with 2AC+, 2A–, 2C–, 1A+, and 1B+ placing high weights on distinct subsets of shared factors (**Fig. 2c**). The other drug groups (2A+, 2C+, 1A–, and 1B–) appear more similar in their weights to the saline control weights (**Fig. 2c**). To quantify how much these behavioral summaries differ across drug groups, we trained a multiclass logistic regression to predict which drug a mouse was given. The classifier achieved 45.8% accuracy on held-out mice. The confusion matrix provides further evidence that some drug groups are much easier to distinguish than others (**Fig. 2d**). Compared to the baseline of 10% accuracy from random guessing, these results indicate that the mouse weights, which summarize syllable usage across space and time, capture relevant differences between groups with most information coming from syllables but the best model-based predictions come from the combination of all three modalities (**Extended Data Fig. 4e**).

The three most informative factors for drug discrimination best described the 1B+, 1A+, and 2AC+ drug groups, respectively. The first factor, with high weights applied to mice in the 1B+ treatment group, emphasizes locomotion syllables near the edge of the arena that decreased in usage over the four major epochs—a finding

that follows from the scalar analyses presented in **Fig. 1**. The second factor, utilized by 1A+ and 2A– mice, describes a shift from pausing behavior to a single slow walking syllable over the course of the behavioral session. The third factor is predominantly featured in 2A+, 2AC+, and 2C– treated individuals and comprises some of the same locomotory syllables as in factor 1, but also a set of epoch-1-associated, wall-oriented rearing behavior as well as locomotion in the center of the arena. Combining the shared factors with the weights of the multiclass logistic regression shows which aspects of behavior are most predictive of each drug group (**Extended Data Fig. 4f**).

The behavior model decomposes spontaneous mouse behavior into shared factors, each encoding a distribution over what, when, and where syllables are used. Each mouse uses a unique mixture of factors, as encoded in their weights. The weights offer a concise, low-dimensional summary that captures core elements of spontaneous behavior but also highlights aspects of drug-affected behaviors that differ among treatment groups. This reduced dimensionality is useful for correlating these data with the Fos maps described below.

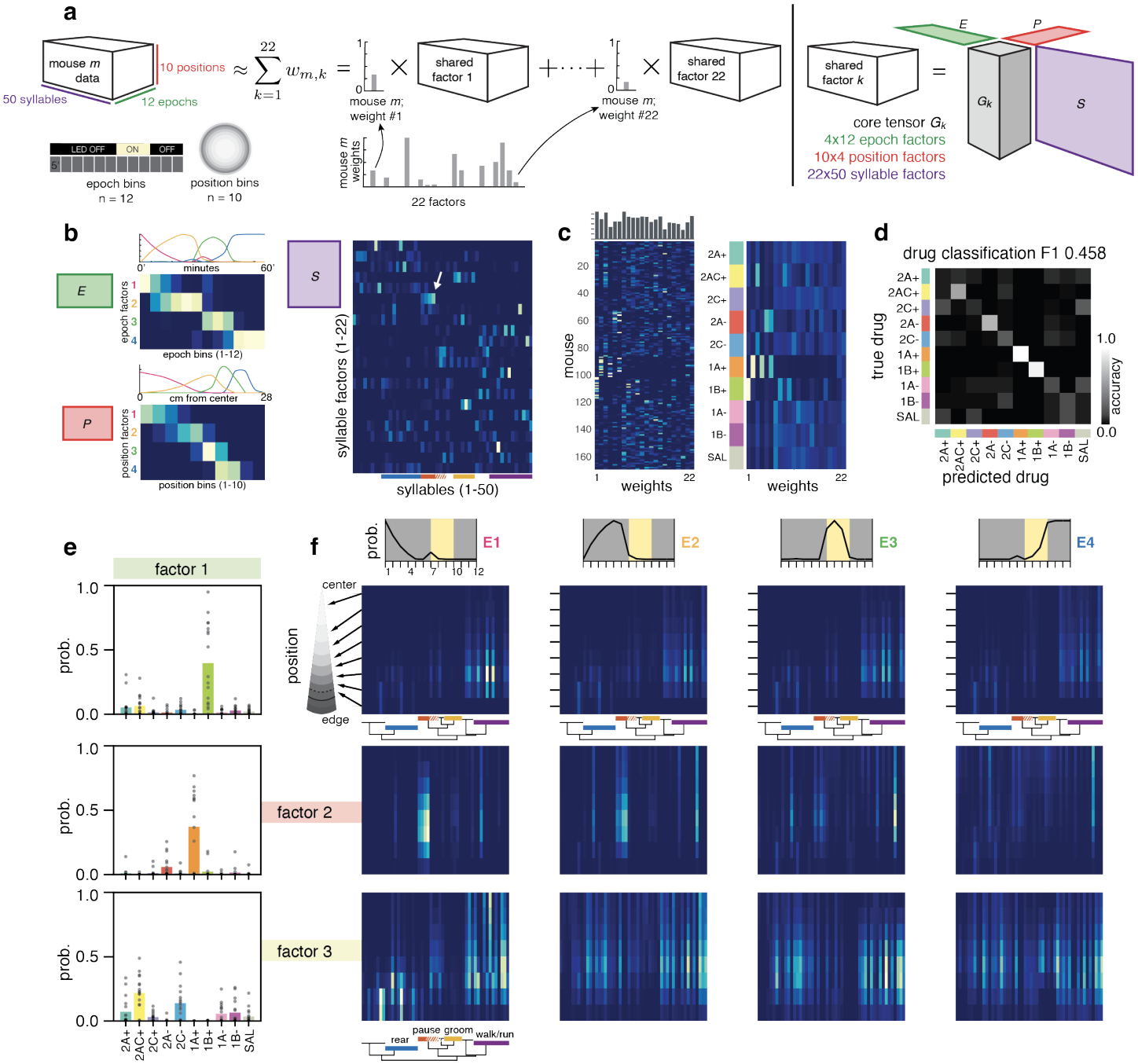


Fig. 2 | Decomposition of behavior. a, Graphical representation of the tensor-decomposition wherein each mouse's behavioral recording is represented as a 3D tensor of frame counts for each combination of syllable (1–50), epoch (1–12; consecutive 5-min bins), and position (1–10; equal area rings, 1 being the center and 10 the periphery of the arena). The resultant factorization includes a shared core tensor *G*, one set of weights each for 4 epoch factors, 4 position factors, 22 syllable factors, and 22 mouse weights. **b,** Top left, the 4 epoch factors arrange into a cascade across the 12 five-min bins from start to end of the recording session. Bottom left, the 4 position factors arrange into a cascade across the 10 concentric position bins of equal area from the center to the edge/wall of the arena. Right, the 22 syllable factors capture subsets of syllables with correlated usage. Subset of pausing syllables in factor 6 is indicated, white arrow. Color code along the lower axis represents rearing, pausing, grooming, and locomotion syllables as established in **Fig. 1**. **c,** The per-mouse weights for each individual (left) and average for each drug group (right). **d,** Confusion matrix of drug classification based on tensor decomposition of behavior. The average accuracy is 0.458, well above the baseline prediction is 0.1 given 10 groups. **e.** The weights on the first three factors differ substantially across drug groups. **f,** Factors are distributions over what, when, and where syllables are used throughout the session. For example, factor 1 shows decreasing levels of locomotion, factor 2 shows elevated pausing, later transitioning to a specific form of locomotion, and factor 3 shows a transition from rearing and walking/running near the periphery to more diffuse behavior by the end of the session.

Effects of drugs on whole-brain Fos intensity

To determine the effect of serotonin receptor agonists and antagonists on neuronal activity, we used whole-brain Fos immunoreactivity as an approximate measure of integrated activity over the period when the mice were exploring the open field under each drug's influence. While sacrificing the temporal dynamics of neuronal activity recording, this approach provides spatial resolution of individual neurons at the scale of nearly the entire brain. Our pipeline (**Fig. 3a–c**) identified on average 7.3 million nuclei (**Fig. 3d**) with continuous Fos intensity levels in addition to binary ON/OFF counts used by most studies. Fos counts and intensity are positively correlated, but with considerable variation along both axes (**Fig. 3e; Methods**). The Fos maps were all registered to the Allen Brain Atlas[38], allowing averaging within brain regions. **Fig. 3f** shows a representative coronal slice of Fos intensity above (red) or below (blue) those in the saline control (presented as raw Fos intensity map in **Fig. 3f$_{10}$**) for 9 drug groups (**Fig. 3f$_{1–9}$**). Each Fos map is an average of ~17 mice, with individual mice exhibiting different degrees of variation in different drug/brain region combinations (**Fig. 3g, h; Extended Data Fig. 5**). The effect of the drugs on the Fos intensity along entire anterior-posterior axis of the brain can be viewed in **Movie S1**.

We use **Fig. 3f** to illustrate a comparison between agonists and antagonists against the same receptor (see **Table S2** for potential explanations of different patterns). (1) The agonist and the antagonist modulate Fos intensity in opposite directions, as is the case for the 2A+ and 2A– in the somatosensory cortex (**Fig. 3f$_1$** and **3f$_2$**, arrows). (2) Only the agonist substantially modulates Fos intensity, as is the case for the 1A+ and 1B+ in bed nucleus of stria terminalis (BNST; **Fig. 3f$_5$** and **3f$_7$**, arrowheads) or the 1B+ in dorsal striatum (**Fig. 3f$_7$**, arrow). (3) Only the antagonist substantially modulates Fos intensity, as is the case for 2A and 2C antagonists in BNST (**Fig. 3f$_2$** and **3f$_4$**, arrowheads). (4) Both the agonist and antagonist modulate Fos intensity in the same direction, as is the case for 2C in somatosensory cortex (**Fig. 3f$_3$** and **3f$_4$**, arrows). (5) Neither the agonist nor the antagonist modulates Fos intensity, as are the cases for the 1A and 2C agonists and antagonists in the dorsal striatum. In general, our Fos maps are consistent with known expression patterns of serotonin receptors[24,26]. For example, 2A is broadly expressed in excitatory neurons of the neocortex[39], consistent with the strongest effect of 2A agonist in cortex. Among the four serotonin receptors examined here, 1B is most prominently expressed in dorsal striatum[24], accounting for the strongest effect of 1B agonist on Fos intensity.

Fig. 3g offers another view of dynamic regulation of Fos intensity across the brain by two of the nine drugs, 2AC+ and 1B+, with quantifications across all mice in 282 Allen regions (**Fig. 3h**; see **Extended Data Fig. 5** for similar data for the rest of the drugs). The rich patterns of Fos intensity across the entire brain and drug conditions (**Fig. 3f–h, Extended Data Fig. 5, Movie S1**) provide a resource for exploring the effect of activating or inhibiting specific serotonin receptors on brain activity. We highlight a small subset of specific findings from our Fos maps in the next section, with a bias towards brain regions in which Fos intensity was modulated by multiple drugs and for which atlas-level summaries inadequately capture the spatial specificity of the modulation.

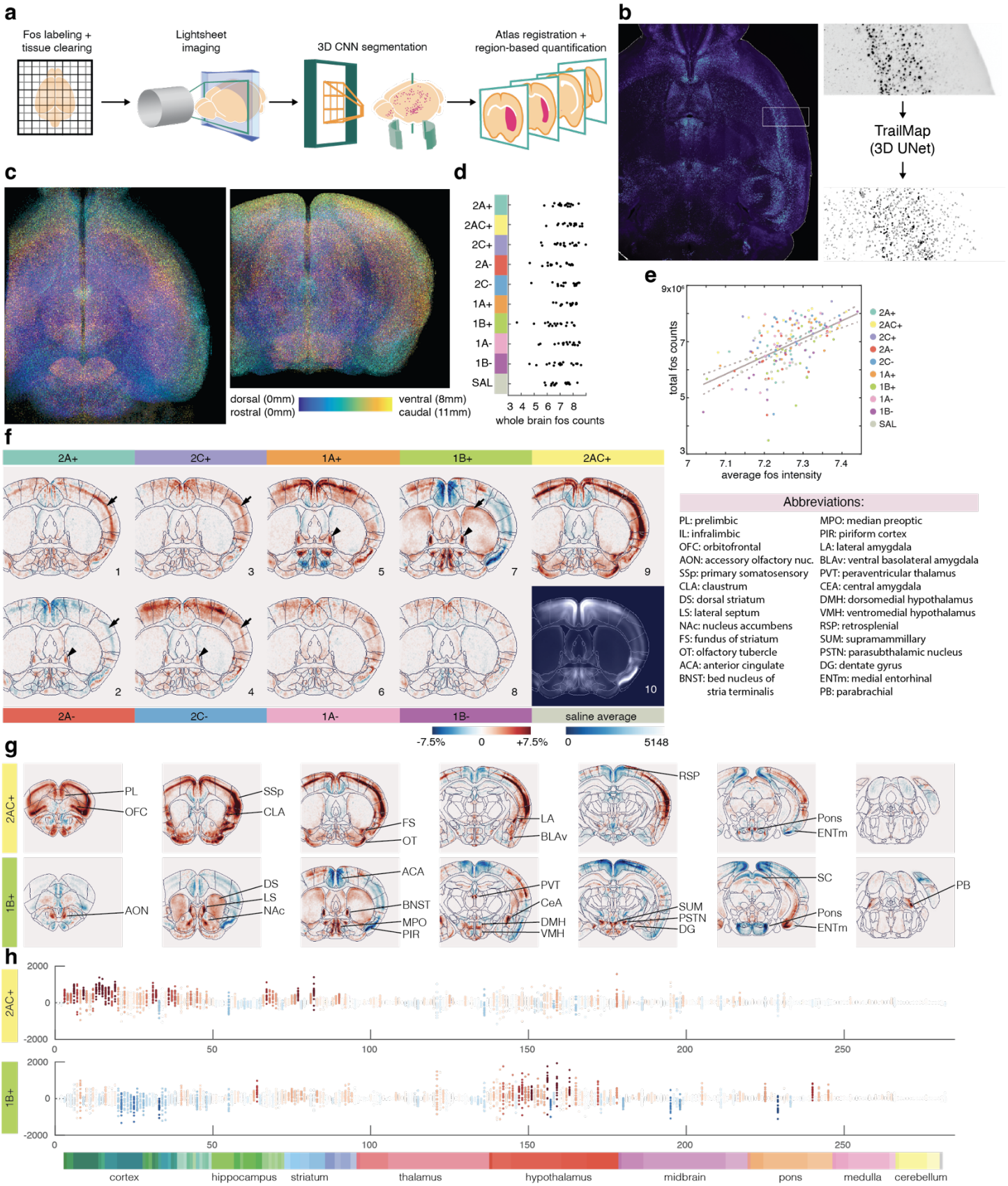


Fig. 3 | Overview of the effect of serotonin receptor agonists and antagonists on whole-brain Fos maps. a, Experimental design for whole-brain analysis pipeline of Fos image data. **b,** Left, horizontal view of a 500-µm Z-projection of raw Fos image data. Right, zoom in of the boxed region from the left image (top), which is segmented by 3D-Unet based TrailMap. Centroids are established as local maxima detected from the probability maps and a 6-connected neighborhood of raw intensity values is averaged to be the intensity of each centroid. **c,** Horizontal (left) and coronal (right) depth-coded views of a single brain's detected

centroids demonstrate even coverage across the entire intact brain. **d,** Total numbers of counted nuclei per mouse and grouped by drug treatment. **e,** Scatter plot of average Fos intensity against the total Fos count (as in **d**) for all 168 mice, color coded by drug treatment. Linear fit shows significant correlation, $R^2 = 0.281$. **f,** A representative 100-µm coronal slice of saline-subtracted Fos intensity maps for each drug treatment group, averaged across individuals within treatment groups. **g,** Coronal slices (as in **f**) across anteroposterior axis for two example drug treatment groups. **h,** Average saline-subtracted Fos intensity for the two drug treatment groups in **g** for all nuclei in each of 282 Allen-brain regions at the bottom, color coded according to Allen Brain Atlas[38]. Each dot represents data from an individual mouse.

Effects of drugs on Fos intensity in select brain regions

The prefrontal cortex, including the orbitofrontal cortex (OFC) and medial prefrontal cortex (mPFC), is strongly innervated by raphe serotonin neurons and in turn provides strong monosynaptic input to raphe serotonin neurons[40–43]. We found that Fos intensity in OFC was up-regulated by 2A+, 1A+, 1B–, but down-regulated by 2A– and 1B+ (**Fig. 4a**). mPFC is traditionally divided into prelimbic and infralimbic cortex (PL and IL). Interestingly, while 1B+ decreased Fos intensity in both regions, 1A+ decreased Fos intensity in PL but increased Fos intensity in IL, whereas 2AC+ increased Fos intensity in PL but decreased Fos intensity in IL (**Fig. 4b**). Thus, serotonin modulates prefrontal cortex activity through multiple receptors with exquisite regional specificity in both magnitude and direction.

Exquisite modulation by multiple serotonin receptors was also evident in many other brain regions. For example, while 2A+ generally increased Fos intensity across the neocortex, it decreased Fos intensity in the retrosplenial cortex. This effect was more pronounced with 2AC+ (**Fig. 4c**). Likewise, in the hypothalamus, while 2C+ increased Fos intensity in both dorsomedial and ventromedial hypothalamic nucleus (DMH and VMH), 2A+ increased Fos intensity in DMH but decreased Fos intensity in VMH, and 1A+ and 1B+ modulated Fos intensity in specific subregions of DMH and VMH (**Fig. 4c**). In the thalamus, 1A+ and 1B+, along with 2A– and 2C–, all increased Fos intensity in a bilaterally symmetric pair of hotspots in paraventricular thalamic nucleus (PVT) (**Fig. 4c**). In the amygdala, 1A+ and 2A– increased Fos intensity in the lateral portion of the central amygdala (CEA) while decreased Fos intensity in the adjacent medial portion of the basolateral amygdala (BLA), whereas 2C– antagonist and 1B+ exhibited a similar effect only on lateral CEA but not on medial BLA (**Fig. 4c**).

The above examples suggest heterogeneity exists within a given brain region in the Allen Brain Atlas as revealed by Fos intensity modulation by serotonin receptor agonists and antagonists. Such sub-region specificity was found in many other brain regions. For example, 1A+, 1B+, and 2AC+ modulated Fos intensity in distinct subregions of the nucleus accumbens (**Fig. 4d**). 1B+ increased Fos intensity highly in the ventral tip of posterior striatum (**Fig. 4e**, arrow). In the ventral hippocampus, 2AC+ increased Fos intensity preferentially in the ventral tip of CA3 (**Fig. 4f**, arrow), whereas 1A+ increased Fos intensity preferentially in the ventral tip of CA1 (**Fig. 4f**, arrowhead). Furthermore, 1A+, 1B+, and 2AC+ modulated Fos intensity in the superior colliculus in a layer-specific manner and in the entorhinal cortex in a subregion-specific manner (**Fig. 4g**). Finally, 1A+ and 2A+, as well as 2A– and 2C–, increased Fos intensity in the ventrolateral subregion of the parabrachial nucleus (**Fig. 4h**), which relay body information to the brain. To our knowledge, specific functions of some of the above subregions have not been described. These findings illustrate the effect of serotonin modulation on highly specific brain regions, consistent with the highly specific collateralization patterns of subpopulations of serotonin neurons[42,44], and suggest that differential serotonin modulation can be used as a means to subdivide brain regions.

Overall, the effects of 2A+ and 2A– were often in opposite directions, with the 2A+ increasing and 2A– decreasing Fos intensity. This is consistent with that fact that 2A is coupled to G_q and thus its activation should depolarize target neurons directly innervated by serotoninergic axons. 1A+ and 1B+ gave some of the largest magnitude of modulation, often with remarkable region specificity (**Fig. 3f, Fig. 4, Extended Data Fig. 5**). While 1A and 1B are both coupled to G_i and thus their activation should hyperpolarize direct target neurons, 1A+ and 1B+ often increase Fos intensity in some brain regions. This could be because they may act indirectly on local inhibitory neurons or axons of long-distance projecting inhibitory neurons innervating the target regions. Curiously, 1A+/1B+ often modulated Fos intensity similarly as 2A–/2C– (e.g.: BNST, **Fig. 3e**; lateral CEA, **Fig. 4c**; parabrachial nucleus; **Fig. 4h**); sometimes 2A– more closely mimicked 1A+ and 2C– mimicked 1B+ (e.g.: medial BLA, **Fig. 4c**), whereas other times it was the other way around (e.g.: OFC, **Fig. 4a**; retrosplenial cortex,

Fig. 4c). These observations suggest cooperation between different combinations of serotonin receptors. We note that the similarities of 1A+ and 2A– effects were also observed in behavioral analyses (**Fig. 2g–i**).

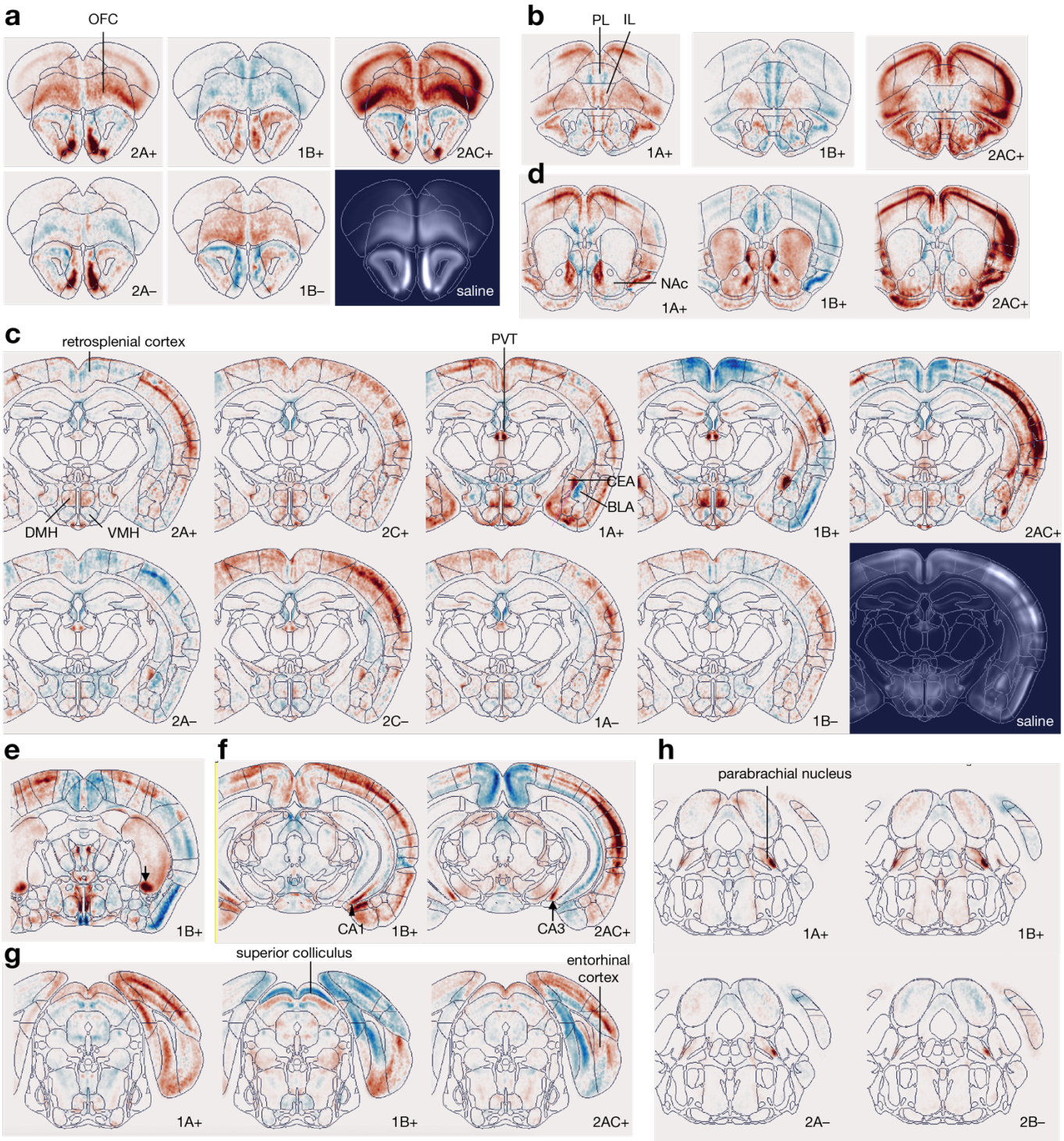


Fig. 4 | Examples of the effects of serotonin receptor agonists and antagonists on Fos maps in selected regions. All images are 100-µm coronal slices highlighting specific brain regions that are modulated by specific subsets of agonists and antagonists as indicated at the bottom right. **a,** Orbitofrontal cortex (OFC). **b,** Medial prefrontal cortex, including prelimic (PL) and infralimic (IL) areas. **c,** A central slice highlighting drug actions on retrosplenial cortex, dorsomedial and ventromedial nuclei of the hypothalamus (DMH and VMH), paraventricular thalamus (PVT), and central and basolateral amygdala (CEA and BLA). **d,** Nucleus accumbens (NAc). **e,** Ventral tip of posterior striatum (arrow). **f,** Ventral tip of hippocampal CA1 (arrowhead) and CA3 (arrow) areas. **g,** Superior colliculus and entorhinal cortex. **h,** Parabrachial nucleus. See text for detailed description of modulation by specific agonists and antagonists.

Decomposition of Fos maps

Fos imaging yields high-dimensional, whole-brain maps for each mouse. We looked for spatial structure in the location and intensity of Fos cells. Such structure could suggest regions that exhibit correlated changes in activity in response to applications of serotonin receptor agonists and antagonists.

To this end, we constructed a model tailored to Fos measurements that can find low-dimensional factors of variation across the collection of mice (**Fig. 5a**). Conceptually, we think of the Fos measurements as a marked spatial point process: each cell has a location (3D) and an intensity (positive scalar). The point process intensity function specifies both the expected number of Fos cells and the expected Fos intensity for each location in the brain. For computational tractability, we approximated the intensity function as piecewise constant within 100-µm voxels. We modeled the intensity function as a weighted combination of nonnegative factors. Each factor is normalized, so that it can be thought of as a distribution over voxels. For each mouse, we jointly modeled the expected Fos cell count and intensity as weighted combinations of these factors (**Fig. 5b, c, Extended Data Fig. 6a; Methods**).

As with the behavioral syllable usage, we found that factor weights describing Fos counts differ substantially between drug groups. For example, weights on factor 1 are largest among mice given 2AC+ (**Fig. 5d**), and factor 1 places highest probability on voxels in the frontal cortical regions (**Fig. 5e**). Factor 11 places higher probability on the striatum and barrel cortex and is most used by mice given 1B+. Factor 7 is most used by mice given 1A+ and places highest probability on the retrosplenial cortex (RSP) and ventral premammillary nucleus (**Fig. 5d, e**).

Factor usage is a strong indicator of drug group, and given the count and intensity weights, we can predict which drug a mouse was given with 51% accuracy using a multiclass logistic regression (**Fig. 5f**). The weights of the logistic regression offer insight into which voxels are most predictive of each drug. **Fig. 5g** shows the combined factors obtained with the logistic regression weights best predicting each drug group. For example, in the depicted coronal slice, increased activity in cortex and decreased midline thalamic activity is most predictive of drug 2AC+, whereas increased activity in caudoventral striatum is most predictive of drug 1B+.

Interestingly, our model reveals hot spots of correlated activity at finer resolution than the Allen atlas affords. By emphasizing the changes filtered by the drug classification task, voxels that are informative for predicting treatment group are highlighted even when the magnitude of their mean saline subtracted intensity is small. In many such cases, the model expects changes in opposite directions for agonist/antagonist pairs even though information on direction of effect was not included in the regression. For example, in the pontine central gray, parabrachial nucleus and locus coeruleus were previously highlighted in **Fig. 3** and **Fig. 4**; however, the factorization also expects changes in Barrington's nucleus (opposing direction of change in 1A+ vs. 1A–) and nucleus incertus (opposing direction of change in 1A+ vs 1A–, and 1B+ vs. 1B–; **Extended Data Fig. 6b**). Amygdalar hot spots in the 2AC+ treated animals are similar to those in the 2A+ condition and a sign inversion in the modeled changes of 2A– treated animals matches in these spots (**Extended Data Fig. 6c**).

Since some of these subregions appear in only a subset of the factors (e.g., CEA, **Fig. 5h**), the maps of specific factors could be considered as being 'coactive' with that subregion. To further dissect this possibility, we used a model-based approximation of the voxel-by-voxel covariance matrix across the entire brain. However, it is infeasible to visualize the entire 350k x 350k covariance matrix that we obtain with 100-µm voxels. Instead, to present a comprehensive summary of brain-wide covariance, we first computed the average covariance for all pairs of Allen atlas regions to obtain the covariance matrix in **Fig. 5i**. Values on the diagonal depict covariance of intra-regional voxels. At this resolution, three large clusters of correlated regions emerged: one cluster consisting of cortical regions; a second cluster of hippocampus, striatum, thalamus, and hypothalamus; and a final cluster of midbrain, pons, medulla, and cerebellum. While some of these relationships are expected (e.g., motor cortices preferentially correlating with striatum and frontal cortices preferentially correlating with amygdalar regions), long-range and less-tested correlations will require experimentation to validate (e.g., pons with subiculum).

Finally, to visualize fine-grained covariance structure at the level of individual voxels, we examined individual rows of the full, 350k x 350k covariance matrix. Each row can be reshaped into a 3D map that illustrates the covariance between all voxels and a single "seed" voxel (**Fig. 5j**). Selecting a single source in dorsal BNST, we see that this voxel is highly correlated with its neighbors, as expected from the hot spot in the corresponding

factors. Less apparent from the factors is that this voxel is highly correlated with activity in the dorsal striatum, PVT, and a hotspot in the central amygdala, and it is anticorrelated with activity in somatosensory cortex. Given the known functional similarity between BNST and central amygdala[45], this finding suggests a methodological approach for hypothesizing relationships between other less well-studied regions. We have developed an interactive viewer to explore brain-wide, 100-µm resolution correlations like these, and several examples are shown in **Extended Data Fig. 6d**.

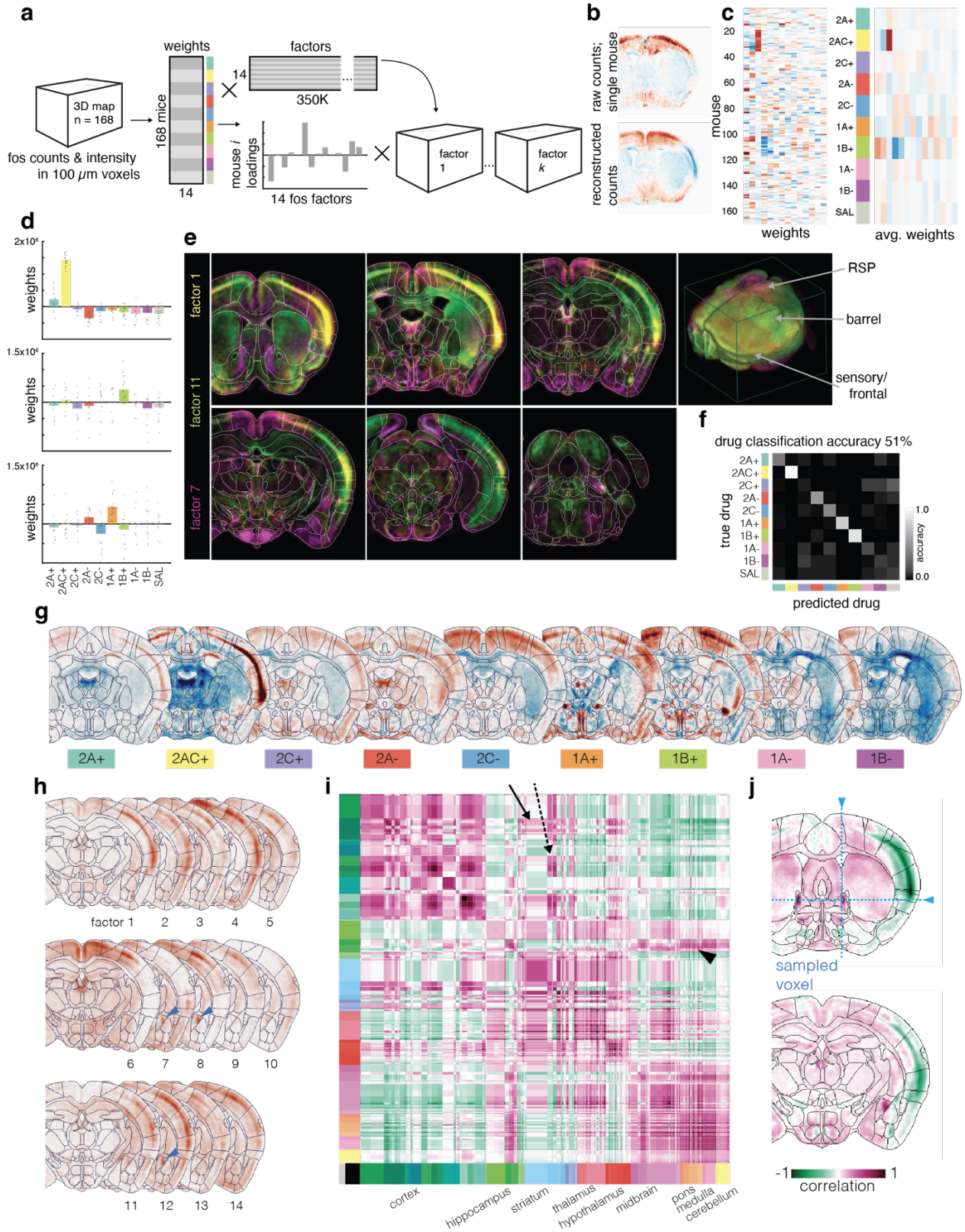


Fig. 5 | Decomposition of Fos maps. a, Schematic representation of the semi-NMF factorization of Fos intensity and count data. All raw data brain maps are factorized into per-mouse weights and 14 separate non-negative Fos factor maps. **b,** Example coronal slice from a saline-subtracted Fos map for a single mouse and the model reconstruction. **c,** The Fos count weights for each mouse (left) and the average for each drug group (right). **d,** The weights for three example factors show differential usage across drug groups. **e,** 2D and 3D visualizations of these three factors highlight regions of overlap and complementarity. **f,** Confusion matrix of drug classification based on weights from the Fos model. The average accuracy is 0.51, well above the baseline prediction of 0.1 given 10 groups. **g,** Weights from the logistic regression combined with the factors of the Fos model show which changes in Fos activity relative to saline are most predictive of each drug. **h,** Single coronal slice from all 14 factors. Arrowhead indicates a region of interest in central amygdala which is apparent in 3 separate factors. **i,** Brain-wide covariance map (as estimated by the 14-dimensional factor model) summarized at the level of Allen Brain Atlas regions. Regions are scaled proportionally to their volume and entries on the diagonal represent the mean covariance of the voxels within that region. Solid arrow, motor cortex and striatum; dashed arrow, frontal cortex and amygdala; arrowhead, subiculum and pons. **j,** A single row of the brain-wide covariance matrix from a seed voxel in BNST highlights the regions and networks that may be coregulated with BNST, including CEA.

Joint neuro-behavioral modeling and prediction

Thus far, we have described the effects of activating or inhibiting specific serotonin receptors on animal behavior (**Fig. 1** and **Fig. 2**) and separately on integrated neuronal activity as measured by whole-brain Fos intensity (**Fig. 3–5**). Given that we have paired behavioral and Fos data for each of the 168 animals, we next examined the relationship between behavior and brain activity, leveraging animal-to-animal variability in both behavior and Fos intensity.

First, we looked for correlations between the per-mouse weights from our behavioral factor model (**Fig. 2**) and our Fos factor model (**Fig. 5**). We used canonical correlation analysis (CCA) to identify pairs of directions in behavioral and neural weight space that exhibited the strongest correlation. We found that the top three canonical directions had correlation coefficients of 0.82, 0.70, and 0.66 respectively (**Extended Data Fig. 7b**), suggesting a strong relationship between the weights of our two models. The CCA directions captured relationships between neural activity and behavior: for example, the first pair of canonical directions captured a correlation between rearing near the periphery of the arena and activity in motor cortex (**Extended Data Fig. 7a**).

Can we predict neural activity from behavior, or vice versa? To do so, we fit a linear model to predict the weights of our behavior factor model given the weights of the Fos factor model (**Fig. 6a**). We trained the linear model on the weights from 75% of the mice and evaluated the predictions on the remaining 25% (**Methods**). With the Fos weights, the linear model can explain a greater fraction of deviance than a model that predicts the average behavior weights. The degree of improvement varies across drug groups, with the 2AC+, 2A–, 1A+, and 1B+ groups showing the greatest increases (**Fig. 6b**). We compared against several baseline models to show that the predictive power in the Fos signal is primarily derived from differences in both Fos and behavior across drug groups (**Methods, Extended Data Fig. 7c**). We also trained models to predict Fos weights given behavior weights and found that predictions were less accurate in this direction. Only the 2AC+ and 1B+ groups showed improvements over baseline when given behavior information (**Extended Data Fig. 7d)**. The poor predictive performance for other drug groups suggests that Fos weights contain additional information about variation in neural activity that is not reflected in overt behavior, at least as measured by syllable usage across time and space.

Next, we used the linear models to predict responses to synthetic "perturbations" of behavior or neural activity (**Methods**). Given the limited performance of the linear mapping from behavior to Fos weights, care must be taken when interpreting predictions in this direction. Nevertheless, our model sensibly predicted increased pausing to correspond with increased Fos activity (i.e., more Fos positive cells) throughout cortex but a decreased Fos activity in the striatum (**Extended Data Fig. 7e, f**). Likewise, we could predict changes to Fos map if we increase syllable frequencies for rearing, grooming, and locomotion, with locomotion producing predictions inverse to those of pausing in cortex and striatum (**Extended Data Fig. 7g**).

Finally, we simulated the effect of synthetic perturbations of Fos activity on behavior, the direction in which our linear models showed greater predictive performance. Increased Fos activity in the striatum (**Fig. 6c**, left) predicted a broad increase in locomotion toward the periphery of the arena, at the expense of a decrease in rearing, pausing, and grooming (**Fig. 6c**, middle and right; **Methods**), consistent with results from our behavior-to-Fos-map prediction (**Extended Data Fig. 7e, f**). We also considered "correlated whole-brain modulation" by

selecting a seed voxel and then modulating brain-wide activity up or down depending on the correlation with the seed voxel (as in **Fig. 5j**). For example, choosing a voxel in ventral premammillary (PMv) led to a brain-wide modulation that includes increased activity in central amygdala (CEA) and decreased activity in cortex (**Fig. 6f**, left). This synthetic perturbation is predicted to correspond to a significant increase in pausing behavior (**Fig. 6f**, right). A predicted increase in pausing is shared across many seed voxels in PMv, but also for seeds in basolateral amygdala (BLA), dorsal lateral geniculate nucleus, and nucleus accumbens while predicted increases in rearing, grooming, and locomotion derive from perturbations originating in lateral cortex, ventral anterior cingulate cortex, and dorsal striatum, respectively(**Extended Data Fig. 8a**). We note that this alternative strategy for estimating brain regions from behavior does not predict changes in Fos, but instead maps spatial regions with high impact on specific behaviors.

To estimate confidence intervals of the predicted change in behavior, we used a stratified bootstrap analysis (**Methods**). Our results show, for example, that correlated whole-brain modulations starting with seed voxels in BLA and CEA both predict increased pausing, but predictions based on BLA seed voxels are more confident (**Fig. 6g**). We repeated the analysis for all 350k voxels throughout the brain, creating a resource for mapping confidence in predictions from any seed voxel. As a select example, seed voxels in striatum predicted the largest increases in locomotion in dorsomedial striatum, but the smaller magnitudes of predicted increases in ventrolateral striatum were equally high confidence (**Extended Data Fig. 8b**). Importantly, as these predictions are for the full behavioral tensor, confidence scores can be generated for each syllable-position-epoch prediction individually or averaged across categories of interest (**Extended Data Fig. 8c**).

Currently, brain-wide perturbations like these are beyond our experimental capability, but as our genetic toolkit for sophisticated perturbations of neural activity continues to grow, our neural and behavioral models offer a means of generating hypotheses for how perturbations of one modality will manifest in the other.

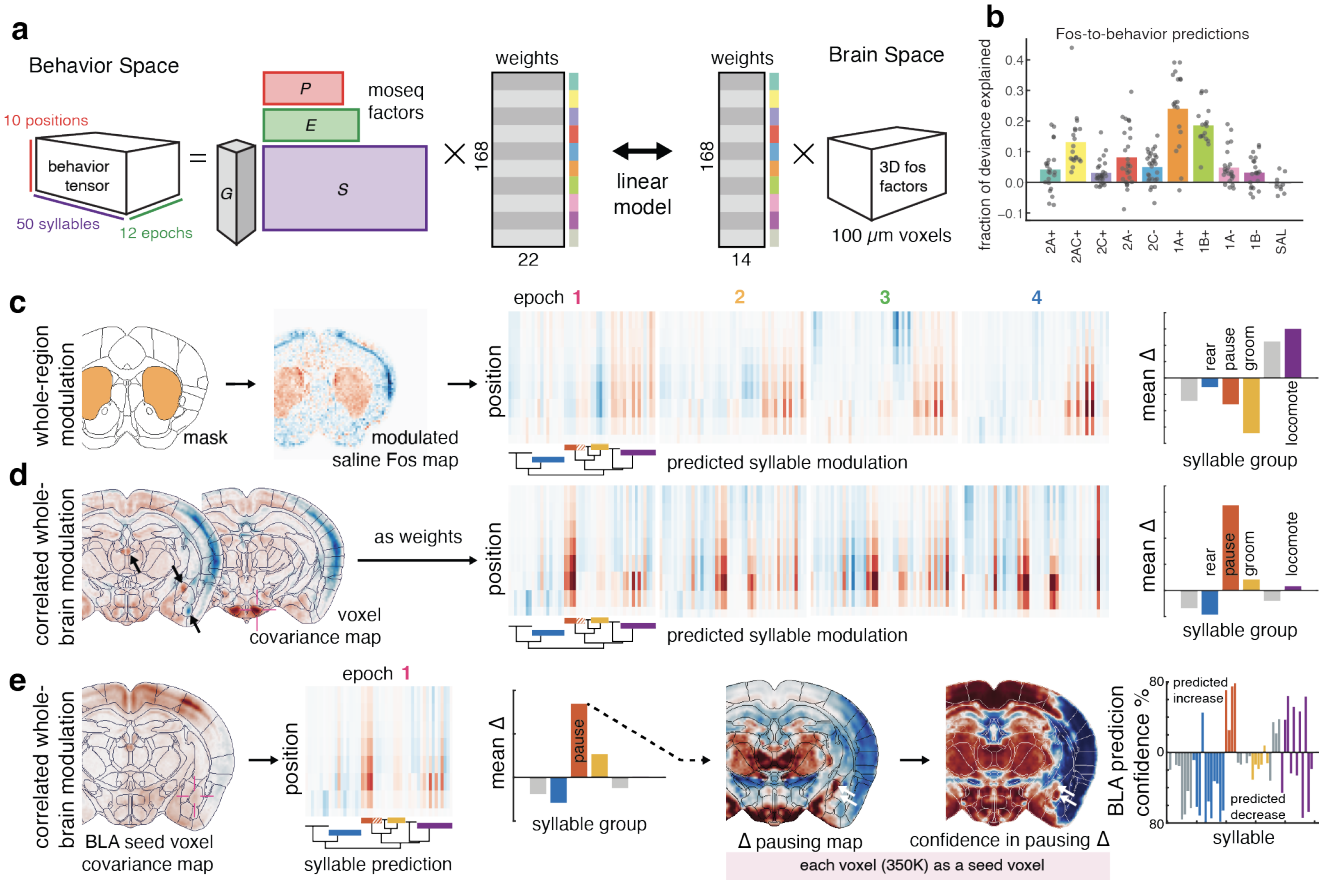


Fig. 6 | Linking behavior and Fos map. a, Schematic of the strategy to link behavior and Fos map using linear regression. **b,** Predictive power of Fos-to-behavior linear model depicted as fraction of deviance explained from held out test mice. **c,** Predicting behavior from synthetically upscaling an average saline brain's Fos activity in the striatum. Following the right-to-left path in **a**

yields predictions of Fos factor weights to estimate the manipulated brain, predictions of the behavior factor weights from the regression, and predictions of the corresponding behavioral tensor, here shown as syllables × position matrix across the 4 major epochs. Bar plot (right) shows the total change in major behavior groups from the unmodulated saline brain's syllable predictions. **d,** Modulating whole-brain patterns using the covariance maps (see **Fig. 5j**) from a seed voxel (here, in ventral premammillary) as positive and negative weights on the entire saline average brain generates behavioral predictions for a brain with increased Fos activity in the seed region, but also modulations in the covarying regions (e.g., PVT, CEA, ventral BLA indicated by arrows). **e,** Bootstrapping model fits to generate confidence maps of Fos-to-behavior predictions. Left, starting with the correlated whole-brain modulation as in **d** with a new seed voxel in BLA, we perturbed the saline average brain with this covariance map and predicted an increase in pausing behavior. As described in **Extended Data Fig. 8a**, the brain-wide map of predicted changes in pausing behavior highlights not only BLA, but also CEA and in fact, selecting CEA as a seed predicts a larger increase in pausing than selecting BLA (white arrows in the " Δ pausing map"). Generating behavioral predictions for all bootstrap model fits and all possible seed voxels creates a confidence map for these predictions. Right, voxel values in this confidence map range from blue (100% of bootstrap fits at this seed voxel show a decrease in pausing) to white (predictions are 50:50 increasing:decreasing pausing) to red (100% of fits predict an increase in pausing). Seeds in BLA have a higher confidence in predicting relatively smaller pausing increases as compared to CEA (white arrows in the "confidence in Δ pausing map"). Far right, averaging confidences for a small hotspot in BLA (27 voxel cube, testing the prediction of all 50 syllables individually at the edge of arena in the time window before LED) shows high confidence in an increase in pausing, high confidence in a reduction of rearing, and mixed predictions for locomotion.

DISCUSSION

Here we report the effect of agonists and antagonists of four widely expressed serotonin receptors on spontaneous mouse behavior, as well as on Fos intensity as a proxy of integrated activity at the scale of the entire mouse brain. We used custom factor analyses, which were tailored to the high-dimensional Fos and behavioral measurements. These models yielded lower dimensional representations of the complex data that shed light on the dominant factors of variation in each domain, capture predictable differences across drug groups, and enabled predictions of behavioral changes following perturbations of the Fos maps and vice versa.

Agonists and antagonists of specific serotonin receptors have been extensively tested in rodents, usually on behaviors aimed for modeling human psychiatric disorders such as anxiety and depression, as selective serotonin reupdate inhibitors (SSRIs) are the most widely prescribed medications for treating anxiety and depression[46,47]. To our knowledge, none of the drugs we used have been examined on spontaneous behaviors using modern analysis methods such as MoSeq, which was demonstrated to outperform traditional scalar measures to distinguish the effect of different drugs[35]. Furthermore, with the exception of 2AC+, we chose our agonists and antagonists to be minimally cross-interfering with non-intended receptors[27–34]. Thus, our large-scale parallel studies of 9 different drugs using both scalar and MoSeq analyses provide a valuable resource on the effect on animal behavior in response to preferentially activating or inhibiting four widely expressed serotonin receptors.

Recent technological advances in imaging and whole-mount histology have enabled examination of Fos (or products of other immediate early genes) at the scale of the entire mouse brain[48–53]. We integrated these methods with deep learning–based image extraction[54], enabling us to identify not only an average of 7 million Fos+ cells but also determine their intensity levels. The comprehensiveness of whole-brain Fos map changes in response to specific serotonin receptor agonists and antagonists enabled many new biological discoveries (see **Results** for details) and provided a rich resource for the field (all data are freely available to the research community). We note that the effects of activating or inhibiting serotonin receptors on Fos maps we report here correspond to integrated activity of mice performing spontaneous behavior in an open arena, including a period with LED light on for heightening anxiety. The effects of manipulating serotonin receptor activity (or other manipulations) on Fos maps corresponding to mice performing other behaviors can be similarly obtained using our data acquisition and analysis pipelines (**Fig. 3**).

Our factor analyses of behavior using tensor decomposition enabled us to not only include MoSeq syllables, but also the position and the temporal epoch in which these syllables were performed. Likewise, our factor analysis of Fos map using a marked spatial point process model enabled us to predict drugs based on Fos maps and which brain regions are most predictive of specific drugs. Furthermore, the factor model yielded an estimate of the brain-wide covariance map, which captures how changes in Fos intensity in one voxel correlates

with changes throughout the rest of the brain. The diversity of serotonin drugs helped drive a range of neural activity patterns to probe this covariance structure. While these maps are specific to this dataset, adding new samples or external datasets could further expand upon or otherwise refine the patterns already present. Among other utility, this covariance map aids in prediction of behavioral changes due to changes in Fos maps, as we discuss next.

Low-rank factor analyses of spontaneous behavior and Fos maps enabled us to explore their relationship. Neural encoding and decoding models have shown impressive abilities to relate real time behavior and neural activity[55–57], but it is unclear *a priori* whether behavior measured in real time with temporal resolution of several hundred milliseconds (duration of average MoSeq syllables) can be related to the Fos map, corresponding to a snapshot of integrated activity of tens of minutes. Canonical correlation analysis and predictive modeling revealed a strong correlation between the two modalities and shows that rich information is contained in the static Fos maps. Importantly, when considering both the low dimensional factorizations and the cross-modal predictions, none of the models included drug identity information or optimized for classification performance. Drug prediction accuracy following behavioral decomposition nearly matches the full dataset's predictive power and just 14 Fos factors predict these 9 drugs, all serotonergic, and some with overlapping targets, well above chance. Importantly, the use of 9 different drugs led to greater variation in behavior and Fos activity than occurs naturally in the saline control group, and the increased dynamic range should improve estimates of neural-behavioral relationships.

The mutual predictions between behavioral syllable usage and Fos maps enable researchers to explore the relationship between brain and behavior and generate specific hypothesis for future experimental validation. While many brain regions of known function are placed into larger context in this joint dataset, we are especially interested in the hypotheses that stem from small subregions, including those that do not obey reference atlas borders and those that strongly correlate with each other across long distances. The whole-brain projections of serotonergic neurons and the diverse set of receptors provide a useful system in revealing surprising heterogeneity in downstream target regions and their interactions. Finally, we note that our cross-modality neuro-behavioral modeling strategy could be applied to analyzing relationship between other datasets with different time scales, such as behavior and gene expression profiling (e.g., from the single-cell/nucleus RNA-sequencing), as well as neural activity (real time or integrated Fos map) and gene expression.

METHODS

Mice

All animal procedures followed animal care guidelines approved by Stanford University's Administrative Panel on Laboratory Animal Care (APLAC). Wild-type mice of the C57Bl6/J strain were used for all experiments (n = 190 8-week-old male animals from the Jackson Laboratories). Twenty animals per week arrived in our animal facility, and were housed on a reverse 12-h light/12h-dark cycle for 12–15 days. The afternoon before a cohort's experiment, a single cage of 5 individuals was split into 5 single-housed cages so that the last mouse of the group wouldn't be in a more anxious state than others before behavioral testing.

Drug preparation

Drug identities and concentrations are listed in **Table 1**. Briefly, stock solutions were made as recommended by the manufacturers in water if possible, but in cases where solubility was an issue, DMSO was used for the stock solution instead. Stock concentrations were diluted in saline (0.9% NaCl) immediately before loading into a syringe at an injection volume of 2 μl/g of mouse weight. The identity of the substance being injected was unknown to the behavioral experimenter. Before intraperitoneal (IP) injection, each mouse was very briefly sedated by isoflurane using a small volume of anesthetic on a nestlet inside an inverted chamber to minimize perceived handling and injection stress. Mice underwent an intraperitoneal injection of one of the serotonin receptor agonists or antagonists, or saline, 30 min prior to behavioral recording. The drugs were counterbalanced so that each drug was present at each recording timepoint during the day-long session and over the days of the week of experimentation. The final drug solutions were always used on the day they were prepared.

Behavioral assay

The camera and arena for behavior were constructed according to specifications listed on https://github.com/dattalab/moseq2-app/wiki. The overhead Kinect camera was frame synchronized via a simulated DAQ with two laterally mounted DMK 33UX273 monochrome cameras (The Imaging Source, data not shown) and a central collimated LED (4900 K, 740 mW, #MNWHL4, #SM2F32-A ThorLabs, powered by a DC4104 driver, ThorLabs) and controlled by MicroManager software.

All behavior was done by one female experimenter between the weekday hours of 10 am and 5 pm. The behavioral room was kept under red light and with white noise present. Each day 5 different animals were tested, and no animal was tested more than once. Each mouse was taken from the reversed light cycle room in a light-tight container and brought to the behavioral room 5 min before IP injection. Behavioral testing began 30 min after IP injection. After this post-injection wait period, the mouse was placed in the center of a round open arena (US Plastic, 14317). The mice were transferred to and from the arena via a 5-inch net pot to minimize handling stress and bedding transfer. In the arena, movement is unrestricted. Video recording was started immediately, and after 30 min, the LED was powered for 15 min, illuminating the center of the arena out to ~65% the radius of the arena or ~40% of the total area. After stimulus, the LED light was turned off and the mice had another 15 minutes of free movement throughout the arena. After cessation of recording, mice were transferred by cup back to their cage and left to rest in the same behavior room for 45 min for Fos expression before perfusion.

After each mouse, the arena was cleaned with 10% Alconox (VWR, 21835-032), water, and 70% ethanol, and left to fully dry before the next mouse was placed in it. One mouse was excluded from further analyses for failure to move in the arena and a second mouse was excluded due to a failure of the LED illumination.

Behavioral analysis

Syllable extraction followed the MoSeq pipeline (https://github.com/dattalab/moseq2-app/wiki/MoSeq2-Extract-Modeling-Notebook-Instructions) (https://doi.org/10.1038/s41593-020-00706-3). Briefly, the 188 hours of recordings were processed to extract an oriented crop of the mouse from each full frame. Scalar information about the *x–y* position of the mouse, the height of the mouse, and the *x–y* area of the animal were also saved for each frame. The cropped data, 80 × 80 pixels × 108000 frames per mouse, were further reduced by principal components analysis (PCA), and subsequently all simultaneously fit by an auto-regressive hidden Markov model. Kappa, the free parameter that regulates the duration of individual states, was selected to fit a model with syllable changepoints (mean = 0.553 sec, mode = 0.233 sec) that best reflect the principal components changepoints (mean = 0.609 sec, mode = 0.200 sec). The fitting procedure yielded 52 syllables to explain 99% of the variance, which we truncated to 50.

For constructing the behavioral tensor, each vector of syllables over time is broken into bins of 9000 frames (5 min), and subsequently a third dimension of position bins using the *x–y* position of the mouse from extracted scalars. Position bins were constructed as ten concentric rings of equal area that capture the most frequently accessed portions of the video frame. In practical terms, the first 6 of these bins represent the floor of the arena, bins 7 and 8 capture a mouse's rearing or exploratory activity against the wall, and the infrequently accessed bins 9 and 10 capture jumps and escape attempts. Crowd movies of each syllable were randomized and blinded and described by 2–4 experts for generating the labels of rear, pause, groom, and walk/run.

Whole-brain Fos immunostaining with iDISCO+

After each mouse's 45-min Fos-expression window following the behavioral recording session, mice were anesthetized with Avertin and transcardially perfused with 20 ml of phosphate-buffered saline (PBS, Invitrogen AM9625) containing 1:1000 10 μg/ml heparin and subsequently 20 ml of ice-cold 4% paraformaldehyde (PFA). All harvested samples were post-fixed overnight at 4°C in 4% PFA in PBS, washed 3x in PBS the following morning and stored at 4°C for 4–7 days. Cohorts of 20 animals from one week of behavioral recordings were processed in parallel with the AdipoClear immunolabeling protocol, as detailed previously[54]. Samples were stained with 1:500 rabbit anti-Fos (Synaptic Sysytems, 226 003) and donkey anti-rabbit AlexaFluor 647 (ThermoFisher Scientific). All antibodies were centrifuged for 30 min at 12,000 rpm and 20˚C before

incubation. Primary antibodies for the complete set of 190 animals were batched together and then pre-measured and aliquoted at the start of the data collection to reduce technical variability in staining.

iDISCO+ Imaging

At least 24 hours after clearing, iDISCO+ samples were imaged on a light-sheet microscope (Ultramicroscope II, LaVision Biotec) equipped with a sCMOS camera (Andor Neo) and a 2×/0.5 NA objective lens (MVPLAPO 2×) with a 6-mm working distance dipping cap. The imaging media was dibenzyl ether. Version v351 of the Imspector Microscope controller software was used. We imaged using 488-nm and 640-nm lasers. The samples were mounted horizontally, illuminated with a single sheet through the right hemisphere cortex, and scanned from dorsal to ventral with a step-size of 3 µm using the continuous light-sheet scanning method with the included contrast adaptive algorithm for the 640-nm channel (20 acquisitions per plane, max lightsheet NA), and without dynamic focusing for the 488-nm autofluorescence. In the *x–y* plane, voxels are 4.0625 µm × 4.0625 µm, yielding a field of view 10.4 mm anterior-posterior × 8.8 mm mediolateral. The small amount of sample shrinkage from the clearing protocol allowed the 6-mm working distance of the objective to capture the entire dorsal-ventral axis of the sample.

Image processing and analysis

Registration to the Allen Reference Atlas CCFv3 and object segmentation were performed as previously described[54]. Briefly, the autofluorescent image acquired with the 488-nm laser was aligned to the reference atlas with a series of linear and non-linear transformations using the Elastix toolbox. These transformation parameters were then used to warp the coordinates of the Fos nuclei.

We used the TrailMap pipeline[54] for segmenting the Fos nuclei from the stained images. While the architecture was unchanged from the previously published axon-detecting scenario, we trained a new model for detecting Fos signal by manually annotating 105 small, cropped volumes of nuclei selected to represent the variance in both signal and noise across the brain and across sample batches. The resultant probability maps were smoothed by convolving with a spherical nucleus-sized kernel and then processed for local maxima detection. Fos intensities for each detected maxima were calculated as the average of the raw image intensity values from a 6-connected neighborhood around the detected centroid.

Batch effects manifested as global shifts in image intensity, both of the detected Fos nuclei and of the background signal in the same channel. By detecting and proportionally scaling this global parameter before segmentation, two batches of 20 animals with reductions in intensity each were aligned with the average batch. However, one batch of 20 animals displayed intensity values too far from the average of the remaining samples and was excluded from further analyses, thus reducing the animals to n = 168.

Tensor decomposition for behavioral data

Let $\mathcal{X} \in \mathbb{N}^{M\times N\times P\times S}$ denote a four-dimensional tensor of non-negative counts $x_{m,n,p,s}$ for each mouse $m = 1, \dots, M$, epoch $n = 1, \dots, N$, position bin $p = 1, \dots, P$, and behavioral syllable $s = 1, \dots, S$. We model the tensor as a weighted combination of factors,

$$E\left[x_{m,n,p,s}\right] = C \sum_{i=1}^{K_M} \sum_{j=1}^{K_N} \sum_{k=1}^{K_P} \sum_{\mathrm{l}=1}^{K_S} g_{i,j,k,\mathrm{l}} \psi_{m,i} \phi_{n,j} \theta_{k,p} \lambda_{\mathrm{l},s},$$

or, more compactly,

$$E[\mathcal{X}] = C \cdot \mathcal{G} \circ \Psi \circ \Phi \circ \Theta \circ \Lambda,$$

where C is a non-negative constant, $\mathcal{G} \in R_+^{K_M\times K_N\times K_P\times K_S}$ is the core tensor, $\Psi \in R_+^{M\times K_M}$ is a matrix with rows ψ_m, $\Phi \in R_+^{N\times K_N}$ is a matrix with rows ϕ_n, $\Theta \in R_+^{P\times K_P}$ is a matrix with columns θ_k, $\Lambda \in R_+^{S\times K_S}$ is a matrix with columns λ_{l}, and $\circ$ denotes a tensor-matrix multiplication. We recognize this as a 4-dimensional Tucker decomposition[36,37]. We impose non-negativity constraints on the factors so that the model can be interpreted as an additive combination of factors. Furthermore, we assume the factors are normalized so that they can be interpreted as distributions over positions, syllables, etc., which we encode via symmetric Dirichlet priors.

We fit the model using *maximum a posteriori* (MAP) estimation. We developed an expectation-maximization (EM) algorithm based on a Poisson data augmentation scheme. After augmentation, each factor admits a closed form coordinate update. The updates are naturally parallelizable. We wrote custom Python software using JAX for compilation and GPU acceleration. The software is available at https://github.com/lindermanlab/dirichlet-tucker. Mathematical derivations and further detail for the EM algorithm can be found in the package documentation.

Hyperparameter selection for tensor decomposition model

The main hyperparameters to be determined are the number of factors, K_M, K_N, K_P, and K_S. We choose these parameters using cross-validation using a random, speckled test set. Specifically, we hold out a random subset of faces $\mathcal{X}_{m,n}$ from the data, i.e., we mask a random subset of (mouse, epoch) pairs. That way, we still have enough observed data to estimate the mouse loadings, ψ_m, for each mouse, and the epoch loadings, ϕ_n, for each epoch.

We evaluate the log likelihood of the held-out data using the parameters estimated from the training data. Rather than reporting the raw log likelihood, we compute the fraction of deviance explained, which is defined as,

$$\text{fraction of deviance explained} = \frac{LL_{\text{model}} - LL_{\text{baseline}}}{LL_{\text{saturated}} - LL_{\text{baseline}}},$$

where LL_{model} is the held-out log likelihood under the Tucker decomposition model, $LL_{baseline}$ is the held-out log likelihood under a baseline of guessing the average behavior, and $LL_{\text{saturated}}$ is the held-out log likelihood under an "oracle" model that predicts that actual test behavior perfectly. The fraction of deviance explained is analogous to the fraction of variance explained in a linear Gaussian model.

We sweep over a four-dimensional grid of numbers of factors $K_M \in \{2,4,\ldots,24\}$, $K_N \in \{2,4,\ldots,8\}$, $K_P \in \{2,4,\ldots,8\}$, and $K_S \in \{2,4,\ldots,24\}$. The bounds of this search space were chosen manually to ensure that higher held out log likelihood could not be achieved with a larger model.

Semi-nonnegative matrix factorization (semiNMF) of Fos data

The Fos data consists of cell counts $\mathcal{C} \in \mathbb{N}^{M\times V}$ and average intensities $\mathcal{Y} \in \mathbb{R}^{M\times V}$ for the M mice and V voxels. We model the cell counts for each mouse and voxel as Poisson random variables, $c_{m,v} \sim \mathrm{Po}(\lambda_{m,v})$, and we model the average intensity as conditionally Gaussian, $y_{m,v} \sim \mathrm{N}(\mu_{m,v}, \sigma_v^2/c_{m,v})$. Since the average intensity is computed from varying number of cells in each mouse and voxel, we rescale the conditional variance accordingly.

We model the mean count and intensity as a weighted combination of factors,

$$\lambda_{m,v} = f\left(\gamma_v + \sum_{k=0}^{K} \beta_{m,k}\, \theta_{k,v}\right)$$

Where γ_v is a per-voxel effect, $\beta_m \in R^{K+1}$ are the per-mouse weights, and $\theta_k \in \Delta^V$ are shared factors. The factors are non-negative and normalized to sum to one, so that they can be viewed as distributions over which voxels are active. The first factor is assumed to be constant, $\theta_0 = \mathbb{1}_V/V$ so that the first weight, β_0, can be interpreted as a per-mouse intercept. The rectifying nonlinear function $f(x) = log(1 + e^x)$ is the "softplus" function, which ensures the expected cell count is non-negative.

The average intensities are modeled analogously,

$$\mu_{m,v} = \gamma'_v + \sum_{k=0}^{K} \beta'_{m,k}\, \theta_{k,v}$$

With their own per-voxel effects, γ'_v, and per-mouse weights, β'_m. The factors, θ_k, are shared with the count model above.

We place a sparsity-inducing Laplace prior with rate α on the per-mouse weights for both counts and average intensity (excluding the per-mouse intercept terms). We place a uniform prior on the shared factors. We place a weakly informative, conjugate, inverse gamma prior on the per-voxel variance σ_v^2.

We fit the model using *maximum a posteriori* (MAP) estimation. We derived a coordinate ascent algorithm that cycles between updating the per-mouse weights, the shared factors, and the per-voxel variances and effects. The Poisson likelihood for counts renders the model non-conjugate. We follow standard practice for sparse generalized linear models (GLMs)[58] and employ a proximal Newton update, which is based on a quadratic approximation of the log likelihood. To account for the sparsity inducing, Laplace prior, the proximal operator employs an inner loop of coordinate ascent to update each weight, $\beta_{m,k}$, one at a time, holding the rest fixed. We use a similar proximal operator to enforce the non-negativity and normalization constraints for the shared factors.

The per-mouse weight updates can be parallelized across mice; likewise, the shared factor updates can be parallelized across voxels. We initialized the model using a heuristic, non-negative singular value decomposition (NNSVD) approach.

We wrote custom Python software using JAX for compilation and GPU acceleration, capitalizing on the parallelizability of the algorithm. The software is available at https://github.com/lindermanlab/fos. Mathematical derivations and further detail on the algorithm can be found in the package documentation.

Hyperparameter selection for Fos semiNMF model

The main hyperparameters to be determined are the number of factors, K, and the sparsity parameter α. We choose these parameters using cross-validation using a random, speckled test set. Specifically, we hold out random, spatially contiguous blocks of voxels for each mouse. That way, we still have enough observed data to estimate the per-mouse weights, β_m and β'_m, and the shared factors, θ_k.

Drug classification

We evaluated how well drug identity could be predicted from the per-mouse weights of the Fos and behavior models by training multiclass logistic regression models to predict the drug labels, $z_m \in \{0,1\}^{10}$ (a one-hot encoding of which drug mouse m was given), given the per-mouse weights (β_m, β'_m) of the Fos model or ψ_m for the behavior model. We used L2 regularization and selected the regularization strength using 5-fold cross validation over a grid of 31 values logarithmically spaced between 1e-15 and 1e15. We standardized each dimension of the regressors before fitting since we are using regularization. We fit the models using Scitkit Learn (sklearn.linear_model.LogisticRegression and sklearn.model_selection.GridSearchCV).

We sorted the factors based on their contribution to drug classification, as measured by the magnitude of the drop in classification accuracy when each factor was left out. The first factor is the one that causes the largest drop in accuracy when omitted from the classifier.

Fos to behavior prediction

After fitting the Tucker decomposition of behavior data and the semiNMF model for Fos data, we trained linear models to predict the per-mouse weights of one model given the weights of the other. Since the per-mouse weights of the behavior model are non-negative and normalized, we clipped and normalized the output of the linear model as follows,

$$\widetilde{\psi_m} = W\beta_m + d$$

$$\widehat{\psi_{m,k}} = \frac{max\{\widetilde{\psi_{m,k}}, 0\}}{\sum_{j=1}^{K_M} max\{\widetilde{\psi_{m,j}}, 0\}}$$

Where $W \in R^{K_M \times K}$ are the weights of the linear model and $d \in R^{K_M}$ is the intercept. We trained the linear model on 4 folds of data, each with a random subset of 75% of the mice, and we evaluated predictions on the held-out 25% for each fold.

We also constructed a "hybrid model" to test if predictions could be improved by incorporating drug identity by incorporating an indicator into the model,

$$\widetilde{\psi_m} = W\beta_m + Dz_m$$

where $D \in R^{K_M \times 10}$ is a matrix of intercepts for each drug, and $z_m \in \{0,1\}^{10}$ is a one-hot indicator of which drug mouse m was given. (The intercept d is redundant in this model formulation.) We train the model using

the true drug labels, and on test data we replace z_m with $\widehat{z_m}$, the predicted probabilities output by the drug classifier trained on Fos weights, β_m.

For end-to-end predictions, we started with the saline average Fos measurements and constructed a synthetic Fos measurement by modulating the Fos cell count, $c_{m,v}$, across voxels to construct a perturbed Fos map. We considered various spatial perturbations, including up-regulating all voxels in a region, or modulating voxels according to their covariance with a seed voxel. We used the semiNMF model with the learned factors, θ_k, to estimate the most likely weights, $\widehat{\beta_m}$, for the perturbed Fos maps. Then we map the estimated weights (and the predicted drugs, $\widehat{z_m}$, for the hybrid model), through the linear model to estimate weights of the behavior model. Finally, we combine the estimated behavior weights with the shared factors of the behavior decomposition model to construct a predicted behavior tensor.

Behavior to Fos prediction

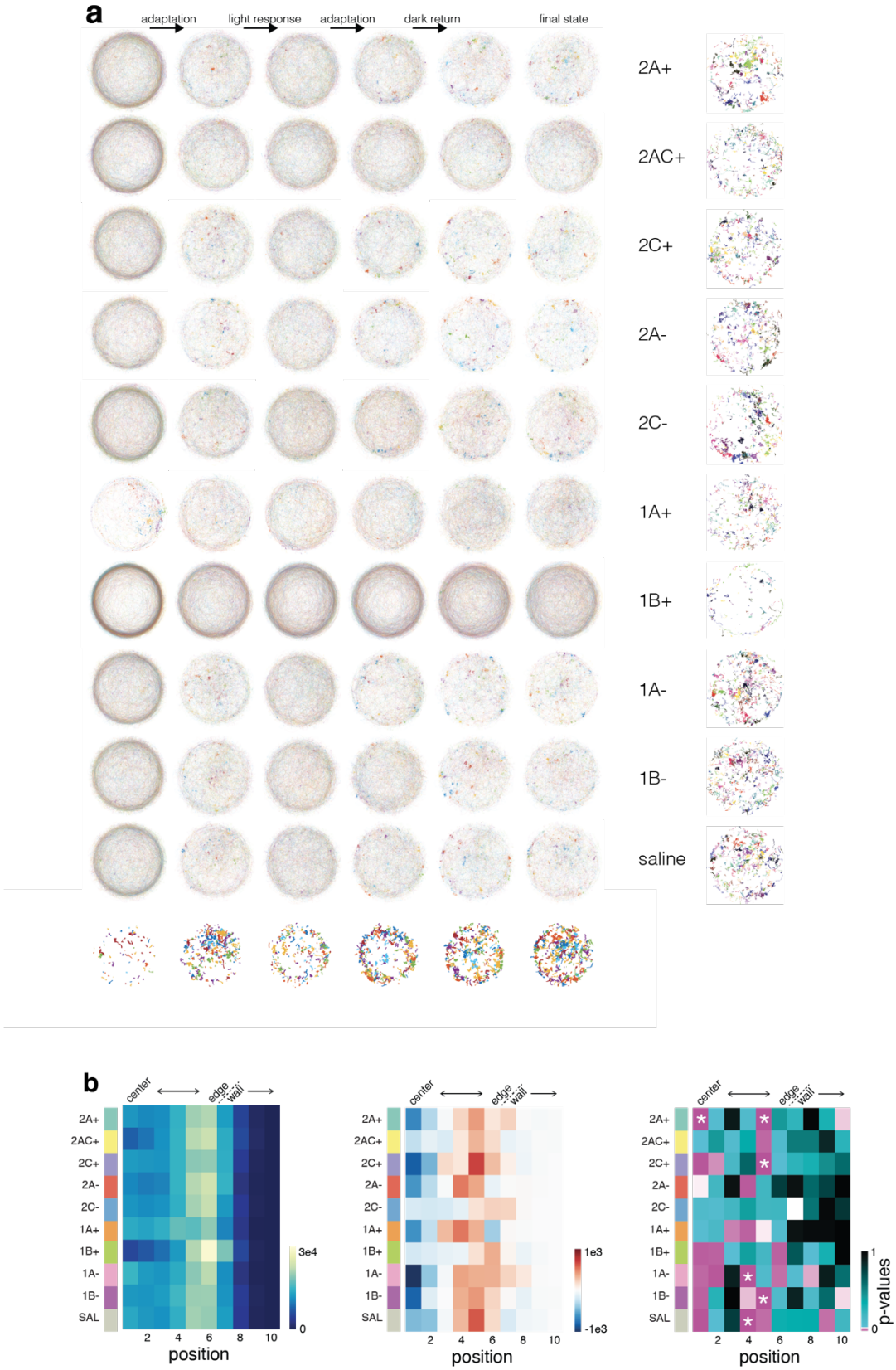
We perform the same analysis in reverse when predicting per-mouse weights β_m of the Fos model given per-mouse weights ψ_m of the behavior model. The only difference is that in this direction, we need to take extra care to account for batch effects in the Fos weights. The Fos data was collected in batches of 20 mice, and each batch differed slightly in the number of cells and average intensity due to differences in the immunostaining, etc. We account for these batch effects by including an extra term in the linear model, Bb_m, where B is a matrix of per-batch intercepts and $\mathrm{b_m}$ is a one-hot indicator of which batch mouse m belongs to.

For the behavior-to-Fos-map prediction, we first constructed a synthetic tensor of syllable counts over space and time by starting with the saline average behavior and then artificially increasing the amount of pausing, grooming, rearing, or locomotion. Note that the total number of video frames per epoch is fixed, so an increase in one category of behavioral syllables necessarily decreases the others. We use the tensor decomposition model to estimate the weights, $\widehat{\psi_m}$, corresponding to the perturbed behavior. Then, using the linear model, we predict the corresponding change in Fos weights. Finally, we map the change in Fos weights from the saline average to a predicted change in Fos measurements.

Bootstrapping analysis

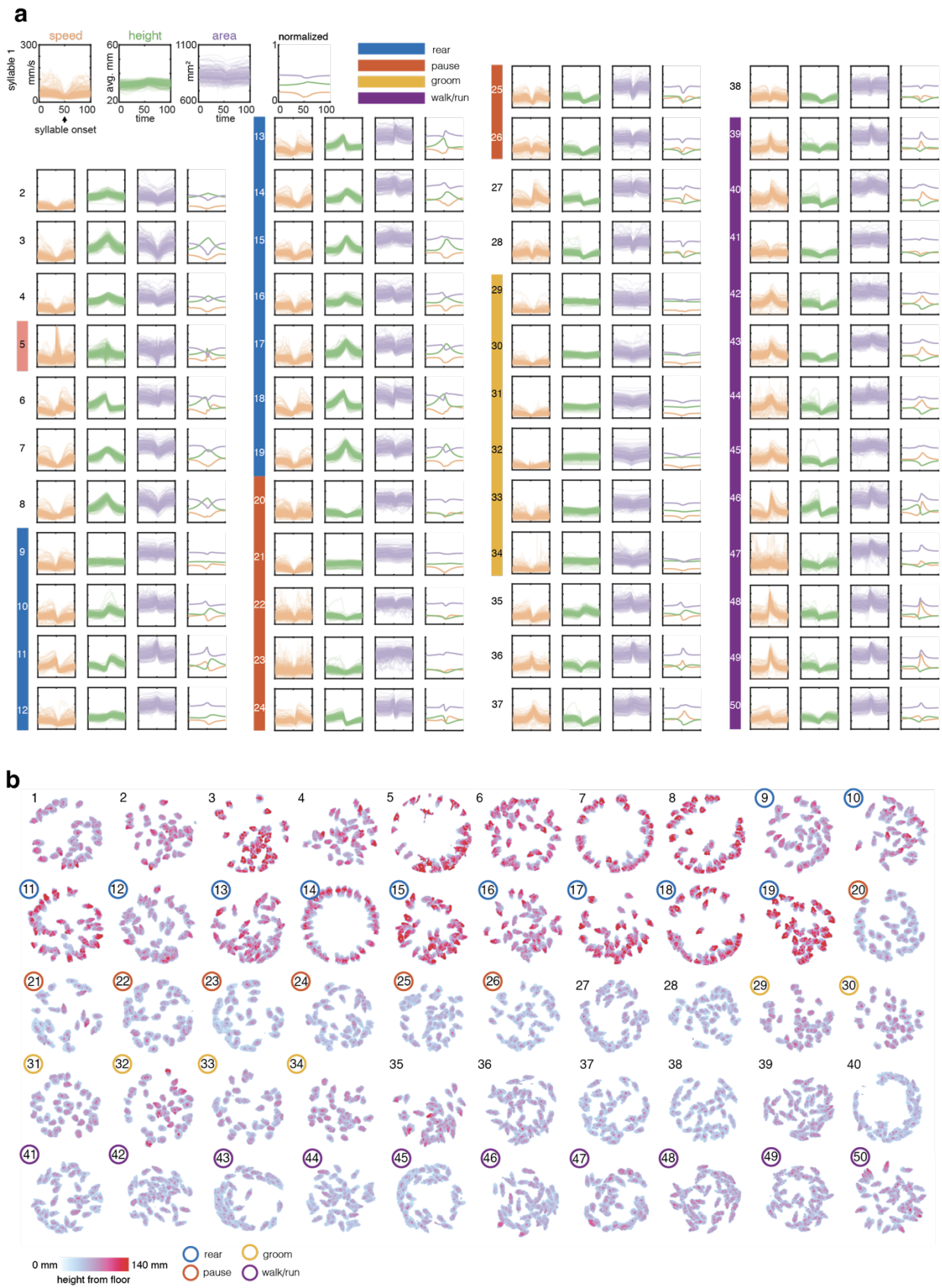
To assess the significance of our Fos to/from behavior analyses we used a bootstrap analysis to estimate confidence intervals of our predictions. We generated 250 replicate datasets using a stratified bootstrap approach: each drug group was resampled with replacement to produce replicate datasets with the same numbers of mice per drug group as in the original dataset. We trained the behavioral tensor decomposition model and Fos semiNMF model on the 250 bootstrap replicates, and we trained linear models to predict Fos weights from behavior weights and vice versa. Then, we computed the distribution of change in behavior for a given change in Fos activity across bootstrap replicates. The quantiles of this distribution provide an estimate of the confidence interval for the prediction.

EXTENDED DATA

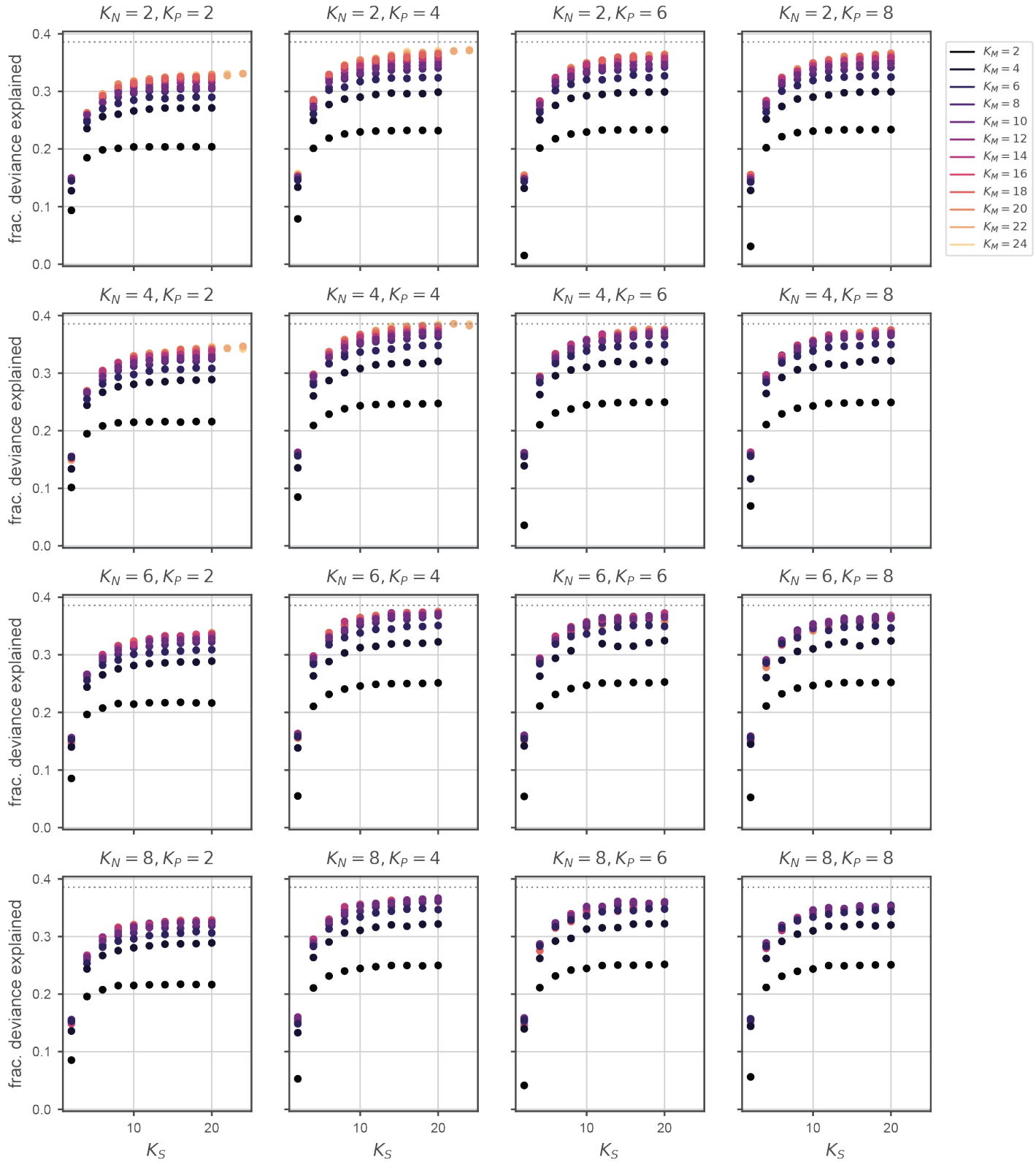


Extended Data Fig.1 | Additional data on the effect of drugs on spontaneous behavior measured by scalar. a, Tracklets for all drug treatment in 6 separate 5-min periods (0–5', 26–30', 31–35', 41–45', 46–50', and 55¨60') from left to right. Right-most column represents slow-speed tracklets in the final 5-min window as described in Fig. 1b. Bottom-most row represents slow-speed tracklets for each of the 6 visualized time bins for saline mice. **b,** Left, drug group average occupancy of 10 spatial bins

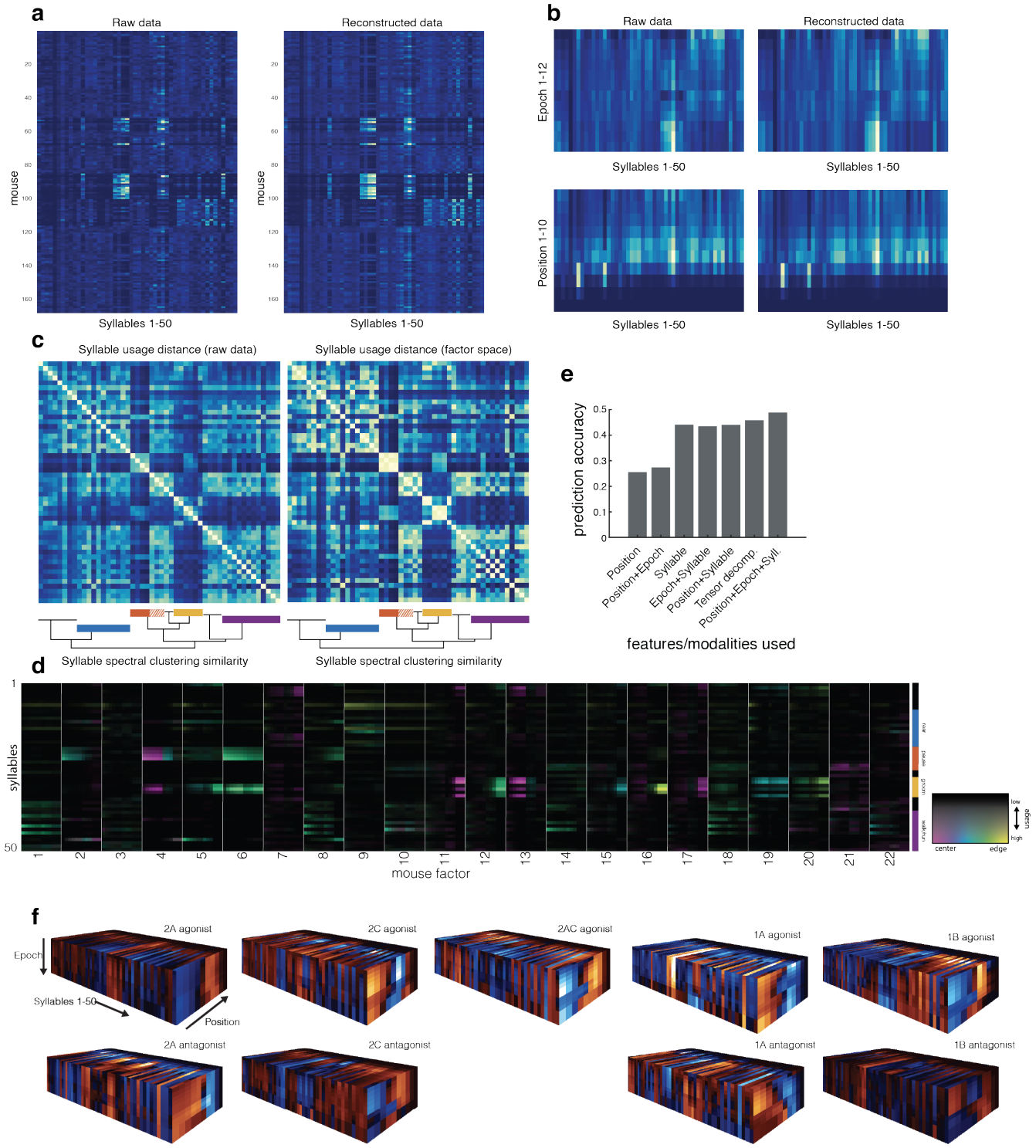
defined as concentric rings of equal area. The edge of the floor is approximately between bins 6 and 7. Mice rearing, jumping, or exploring the wall surfaces appear in bins 7-8 and above. Rows sum to 1.08×10^5, the number of frames in 60'. Center, differential occupancy of position bins in the five minutes before and after the LED onset. Rows sum to zero. Right, t-test for deviation from zero for the LED onset response; asterisks represent p-values passing a Bonferroni-corrected α for m=10. 2AC+, 2C–, and 1A+ treated groups did not show a significant reduction in center occupancy and the 2C– group also did not show an increase in occupancy of the edge of the arena.



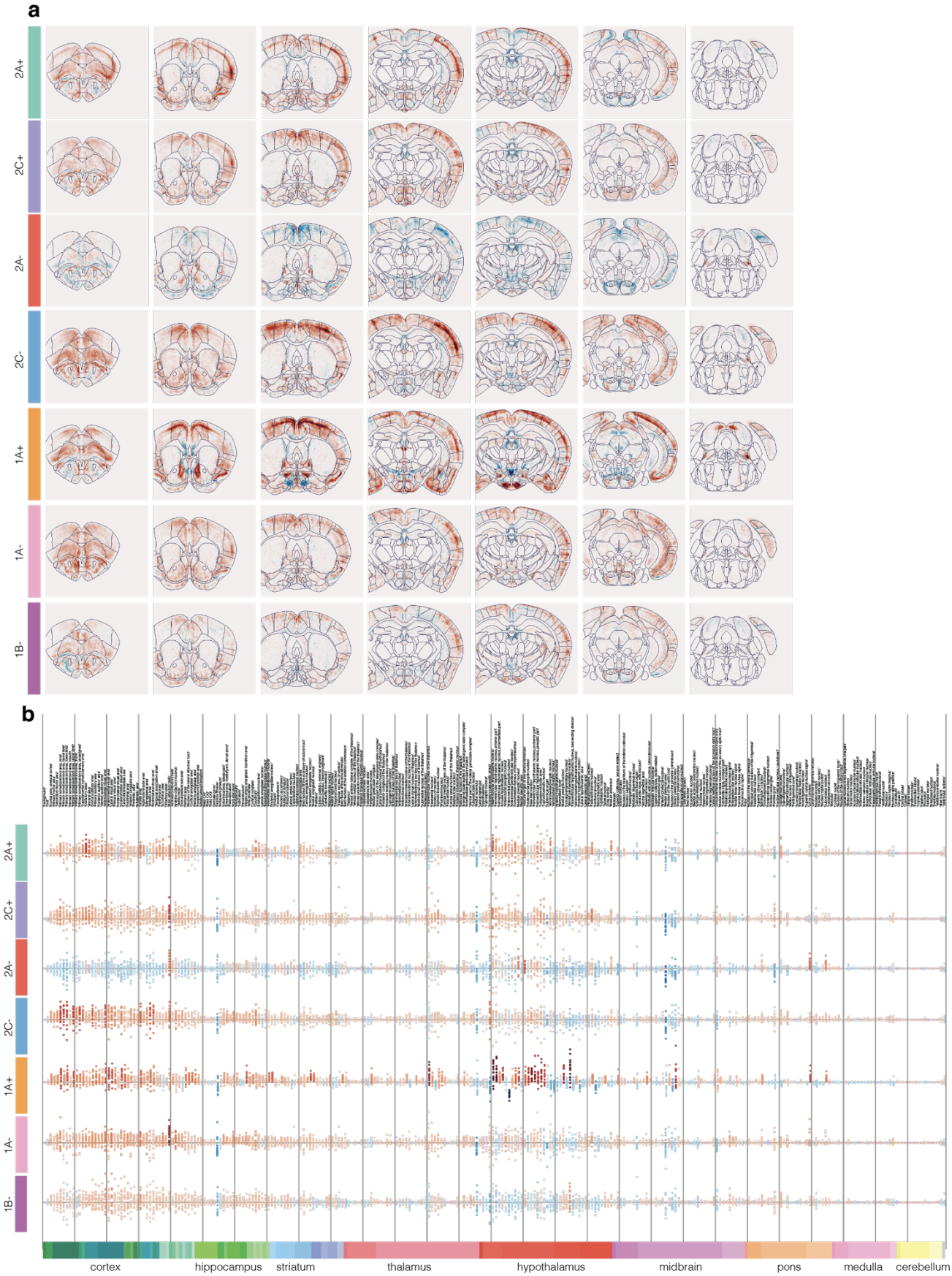
Extended Data Fig. 2 | Additional data on the effect of drugs on spontaneous behavior as measured by MoSeq. a, Speed, height, and area plots for all 50 syllables. Time = 50 represents time-locked onset of the syllable. The four major behavioral clusters are highlights with colored bars. Syllable 5, the jumping syllable, is also highlighted. **b,** Single frame extracted from crowd movies of each individual syllable during the display of the syllable. Height of each individual pixel is color coded from the depth camera output and highlights the mouse body shape and position for syllables with high average height plots in **a**.



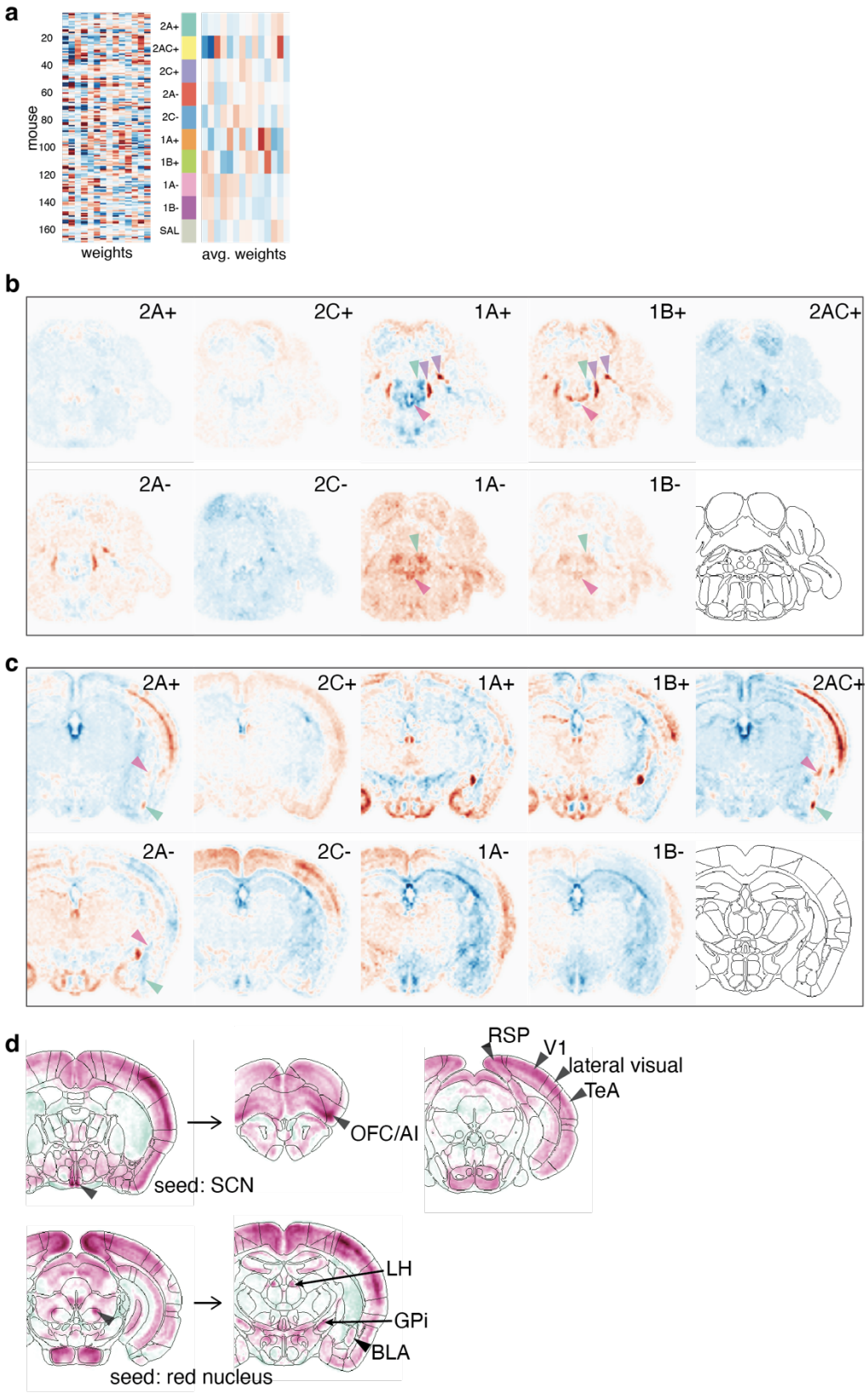
Extended Data Fig. 3 | Hyperparameter selection for behavior model. The hyperparameters include the number of shared factors K_M, the number of epoch factors K_N, the number of position factors K_P, and the number of syllable factors K_S. The optimal setting of hyperparameters was chosen by a grid search, which identified the combination that yielded the highest fraction of deviance explained on held-out test data (**Methods**).



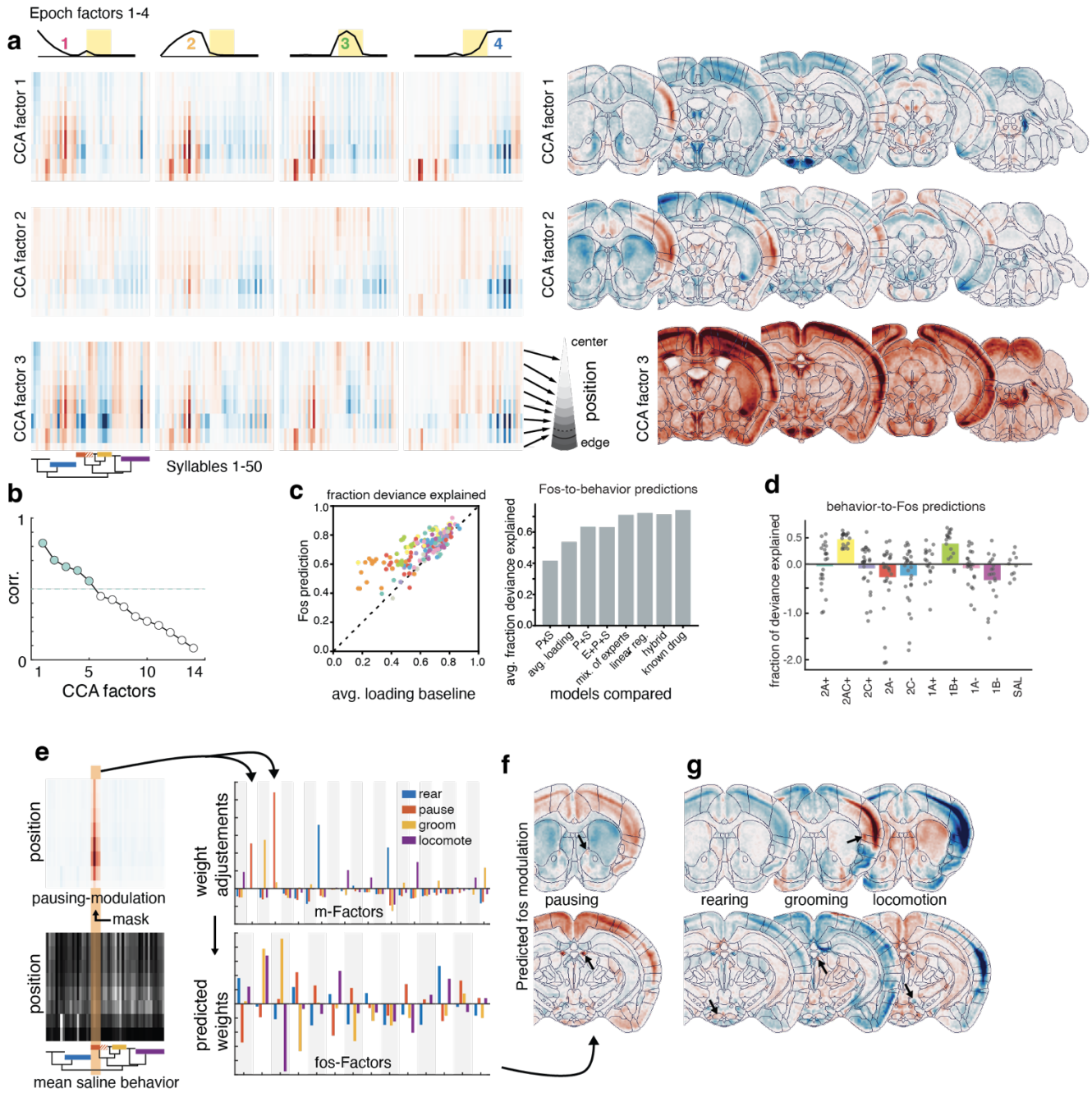
Extended Data Fig. 4 | Decomposition of behavior. a, Side-by-side view of syllable usage by mouse from raw data (left, as in Fig. 1g) and reconstructed data from the Tucker decomposition (right). **b,** Side-by-side view of raw and reconstructed behavioral data from the behavioral tensor. Top, epoch by syllable; bottom, position by syllable. **c**, Cosine distance between individual syllable usage vectors across n = 168 mice is similarly captured by the reconstructed data. **d,** All 22 mouse factors from the decomposition shown as usage heatmaps of 50 individual syllables over 12 epochs. The intensity of the color corresponds to usage while the hue corresponds to the modal position bin (1–12) for that syllable/epoch pairing. **e,** Prediction accuracy for the drug classification multi-class logistic regression, using one or more modalities. For example, "Position" model predicts drug solely based on a mouse's average position in the arena. "Tensor decomp" indicates performance of the logistic regression using the 22-dimensional mouse weights from the Tucker decomposition model, which is much lower-dimensional than the raw, 6000-dimensional position+epoch+syllable count vector. **f,** The weighted combination of shared factors using the weights of the logistic regression model show which epochs, syllables, and positions are most predictive of each drug. Elements of the tensor are summed and represented on each face.



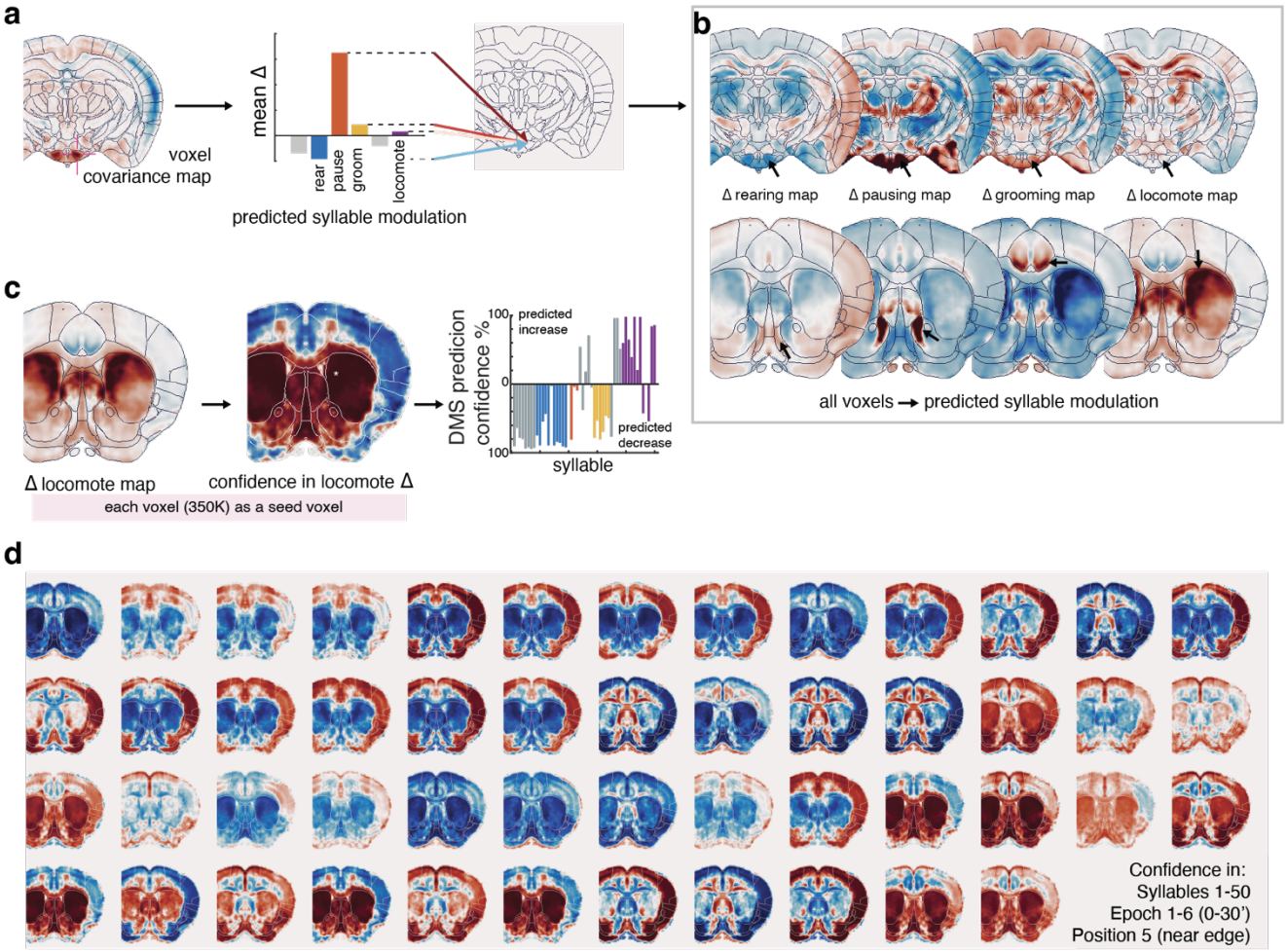
Extended Data Fig. 5 | Additional data on drugs' effect on whole-brain Fos map. a, 100-µm coronal slices (as in **Fig. 3g**) across anteroposterior axis for the 7 remaining drug treatment groups. **b,** Average saline-subtracted Fos intensity (as in **Fig. 3h**) for the 7 remaining drug treatment groups in **a** for all nuclei in each of 282 Allen-brain regions at the bottom, color coded according to Allen Brain Atlas[38]. Each dot represents data from an individual mouse.



Extended Data Fig. 6 | Additional data on decomposition of Fos map. a, The Fos intensity weights for each mouse (left) and the average for each drug group (right). **b,** Caudal slice from prediction filtered factors (as in **Fig. 5g**). Arrowheads mark PBN and LC (purple), Barrington's nucleus (green), and nucleus incertus (pink). **c,** Data as in **b**, a slice highlighting the expected changes in amygdalar hotspots of 2A+ interacting drugs with a sign inversion to the expected change in the antagonist. BLAv, green arrowhead, lateral amygdala, pink arrowhead. **d,** Voxel-wise covariance maps as in **Fig. 5j** for seed voxels in SCN (top) and red nucleus (bottom). Arrowheads mark notable hotspots of high positive correlation.



Extended Data Fig. 7 | Linking behavior and Fos map using canonical correlation analysis (CCA) and linear regression. **a,** CCA generates 14 canonical factors, the first three shown here. Factors are in both behavior space (left, represented as syllable × position heatmaps across the 4 epochs. Values are expected difference from saline) and brain space (right, as expected difference from saline). Note that the top two factors have both positive and negative values, but factor 3 is all positive values. **b,** Correlation values for all 14 canonical factors, the first of which is 82% correlated between the two modalities. The first 5 factors are over 50% correlated (green circles). **c,** Left, individual mouse data summarized in **Fig. 6b** for Fos-to-behavior prediction direction (see **Methods**). Right, fraction of deviance explained in the Fos-to-behavior direction across models selected with varying quantities of information. Linear regression, our selected model, is second only to the inclusion of drug identity. **d,** Predictive power of behavior-to-Fos linear models depicted as fraction of deviance explained from held out test mice. **e,** Predicting Fos maps from synthetic modulations on average saline behaviors. Left, upmodulating pausing syllable usage from the saline average behaviors (gray) yields an artificial tensor that has extra pausing (red). The behavioral factor weights for this artificial pausing tensor are predicted (orange bars only, increased weights on factors 2 and 4), then using the coefficients from the linear regression, the Fos factor weights are predicted and combined with all 14 Fos factors to generate predicted brain maps. **f,** Values represent expected change in Fos from a similarly predicted saline average map without modulation. **g**, Similarly produced maps for rearing, grooming, and locomotion, following the same synthetic strategy of upmodulating groups of syllables across all positions and epochs.



Extended Data Fig. 8 | Linking behavior and Fos map via correlated whole-brain modulation. a, Voxel-wise covariance map for a seed voxel in ventral premammillary nucleus (PMv), as seen in **Fig. 6f**, is used as a perturbation map for predicting changes in behavior. Starting with the average saline brain, 10 artificial cells are added to the seed voxel (average of 23 cells per voxel across all voxels in all brains) and up to 10 artificial cells are added to or subtracted from all other voxels, weighted by each voxel's covariance (–1 to 1) with the seed. Predicting behavior (as described in **Extended Data Fig. 7e** in the behavior-to-Fos direction) generated the barplot to the right. In this example, voxels in the hotspot in PMv (at left) will have high scores for pausing behavior, low positive scores for grooming, and negative scores for rearing. **b,** By generating a brain map wherein these seed voxels contain the values of the predicted changes in behavior, we can visualize the extent of the hotspot in PMv that predicts pausing behavior and other behaviors as well. Repeating this procedure at every seed voxel in the brain yields complete maps of brain regions that predict high magnitude behavioral modulation for each specific subset of behavior. Note that these maps are not imputed Fos values at each voxel; rather, these values represent the predicted change in each behavior that follows from a standardized perturbation originating at each voxel. **c,** Left, the map of all seed voxels whose perturbation predicts modulation of locomotion behavior. Large positive values in dorsomedial striatum indicate that synthetically high Fos counts in this region (and the brain-wide Fos pattern that covaries with it) predict increased locomotion in the arena. Center, after testing each seed voxel's prediction across bootstrapped model fits, we generated a voxel-wise map of the percentage of the fits at each voxel that shows a positive (red) or negative (blue) modulation of locomotion. Values approaching 1.0 indicate high confidence that perturbations at that voxel predict increases in locomotion, but this does not carry information about the magnitude of the predicted change. Right, as confidence scores from each seed voxel can be generated for each syllable+epoch+position triplet, we averaged the confidence scores for predictions of all 50 syllables (in epochs 1–6, position bin 5) following perturbations at each seed in a 27-voxel cube at the asterisk in the previous map. Perturbations in this region confidently predict increases in most locomotion syllables and confidently predict decreases in grooming and rearing syllables. **d,** Confidence maps, as in **c,** for each of the 50 syllables, averaged across epoch 1–6, position bin 5. Values in a given voxel approaching 1.0 and –1.0 indicate high confidence that a perturbation originating at that voxel will well-predict modulation of that specific syllable at the edge of the arena and in the first 30 min.

Table 1: Serotonin receptor agonists and antagonists used in this study

Group[1]	Drug	Target[2]	Dose (mg/kg)
1	25CN-NBOH	2A+	3
2	DOI	2A/2C+	1
3	PF-04479745	2C+	5
4	MDL100907	2A–	1
5	SB-242084	2C–	0.75
6	NLX-101	1A+	0.75
7	CP 94253	1B+	10
8	(S)WAY-100135	1A–	10
9	SB-216641	1B–	5
10	saline	control	none

[1] With the exception of Fig. 3f, we list drugs according to this sequence.
[2] Listed here are the serotonin receptors that have the highest affinity for a particular drug, even though most drugs have lower affinity binding to other receptors.

Movie S1: Flythrough of saline-subtracted Fos intensity maps for 9 drugs and the saline group (bottom right).

Table S1: Description of 50 MoSeq syllables

Table S2: Possible interpretations of agonist and antagonist effects

Fos intensity	Example (in **Fig. 3f**)	Possible interpretations[1]
Enhanced by agonist; suppressed by antagonist[2]	2A in somatosensory cortex	Receptor is expressed and activated in saline control by the open field behavior at an intermediate level
Modulated by agonist; no effect by antagonist	1B in BNST or dorsal striatum (Fig. 3e)	Receptor is expressed but not substantially activated in control by the behavior
No effect by agonist, modulated by antagonist	2A or 2C in BNST	Receptor is activated at a saturation level by the behavior in control, so only downregulation has an effect
Agonist and antagonist have the same effect	2C in somatosensory cortex	Receptor is expressed in mutually connected excitatory and inhibitory neurons in the same region, and agonist and antagonist preferentially act on receptors in either of the populations
Neither agonist nor antagonist has an effect	2A or 2C in dorsal striatum	Receptor is not substantially expressed in the brain region of interest

[1] We list only one of several possible interpretations. One complicating factor is that serotonin can act on receptors expressed on presynaptic terminals of neurons far away from the region of interest.
[2] The opposing effects can be the other way around (enhanced by antagonist, suppressed by agonist). The simplest interpretation is that activation of the receptor suppresses neuronal activity (as is the case for 1A and 1B receptors, which are coupled to G_i and known to suppress neuronal activity). The other interpretation is that the receptor activates local inhibitory neurons, which in turn suppress the activity of nearby neurons.

References

1. Fyhn, M., Molden, S., Witter, M. P., Moser, E. I. & Moser, M.-B. Spatial representation in the entorhinal cortex. *Science* **305**, 1258–1264 (2004).
2. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
3. O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* **34**, 171–5 (1971).
4. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
5. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).
6. Kim, T. H. *et al.* Long-Term Optical Access to an Estimated One Million Neurons in the Live Mouse Cortex. *Cell Rep.* **17**, 3385–3394 (2016).
7. Datta, S. R., Anderson, D. J., Branson, K., Perona, P. & Leifer, A. Computational Neuroethology: A Call to Action. *Neuron* **104**, 11–24 (2019).
8. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**, 1281–1289 (2018).
9. Weinreb, C. *et al.* Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics. *Nat. Methods* **21**, 1329–1339 (2024).
10. Wiltschko, A. B. *et al.* Mapping Sub-Second Structure in Mouse Behavior. *Neuron* **88**, 1121–1135 (2015).
11. Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M. & Keller, P. J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat Methods* **10**, 413–20 (2013).
12. Atanas, A. A. *et al.* Brain-wide representations of behavior spanning multiple timescales and states in C. elegans. *Cell* **186**, 4134-4151.e31 (2023).
13. Chen, X. *et al.* Brain-wide Organization of Neuronal Activity and Convergent Sensorimotor Transformations in Larval Zebrafish. *Neuron* **100**, 876-890.e5 (2018).
14. Kato, S. *et al.* Global brain dynamics embed the motor command sequence of Caenorhabditis elegans. *Cell* **163**, 656–669 (2015).
15. Sheng, M. & Greenberg, M. E. The regulation and function of c-fos and other immediate early genes in the nervous system. *Neuron* **4**, 477–85 (1990).
16. Allen, W. E. *et al.* Thirst-associated preoptic neurons encode an aversive motivational drive. *Science* **357**, 1149–1155 (2017).
17. DeNardo, L. A. *et al.* Temporal evolution of cortical ensembles promoting remote memory retrieval. *Nat Neurosci* **22**, 460–469 (2019).
18. Liu, X. *et al.* Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature* **484**, 381–5 (2012).
19. Reijmers, L. G., Perkins, B. L., Matsuo, N. & Mayford, M. Localization of a stable neural correlate of associative memory. *Science* **317**, 1230–3 (2007).
20. Dahlstrom, A. & Fuxe, K. Localization of monoamines in the lower brain stem. *Experientia* **20**, 398–9 (1964).
21. Steinbusch, H. W. Distribution of serotonin-immunoreactivity in the central nervous system of the rat-cell bodies and terminals. *Neuroscience* **6**, 557–618 (1981).
22. Müller, C. P. & Cunningham, K. A. *Handbook of the Behavioral Neurobiology of Serotonin, 2nd Edition*. (Academic Press, London, 2020).
23. Okaty, B. W., Commons, K. G. & Dymecki, S. M. Embracing diversity in the 5-HT neuronal system. *Nat Rev Neurosci* **20**, 397–424 (2019).

24. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–76 (2007).
25. Nichols, D. E. & Nichols, C. D. Serotonin Receptors. *Chem. Rev.* **108**, 1614–1641 (2008).
26. Vilaró, M. T., Cortés, R., Mengod, G. & Hoyer, D. Distribution of 5-HT receptors in the central nervous system: an update. in *Handbook of Behavioral Neuroscience* vol. 31 121–146 (Elsevier, 2020).
27. Fantegrossi, W. E. *et al.* Hallucinogen-like effects of 2-([2-(4-cyano-2,5-dimethoxyphenyl) ethylamino]methyl)phenol (25CN-NBOH), a novel N-benzylphenethylamine with 100-fold selectivity for 5-HT2A receptors, in mice. *Psychopharmacology (Berl.)* **232**, 1039–1047 (2015).
28. Fletcher, A. *et al.* WAY100135: a novel, selective antagonist at presynaptic and postsynaptic 5-HT1A receptors. *Eur. J. Pharmacol.* **237**, 283–291 (1993).
29. Kehne, J. H. *et al.* Preclinical characterization of the potential of the putative atypical antipsychotic MDL 100,907 as a potent 5-HT2A antagonist with a favorable CNS safety profile. *J. Pharmacol. Exp. Ther.* **277**, 968–981 (1996).
30. Kennett, G. A. *et al.* SB 242084, a selective and brain penetrant 5-HT2C receptor antagonist. *Neuropharmacology* **36**, 609–620 (1997).
31. Koe, B. K., Nielsen, J. A., Macor, J. E. & Heym, J. Biochemical and behavioral studies of the 5-HT $_{1B}$ receptor agonist, CP-94,253. *Drug Dev. Res.* **26**, 241–250 (1992).
32. Newman-Tancredi, A. Biased agonism at serotonin 5-HT $_{1A}$ receptors: preferential postsynaptic activity for improved therapy of CNS disorders. *Neuropsychiatry* **1**, 149–164 (2011).
33. Price, G. W. *et al.* SB-216641 and BRL-15572--compounds to pharmacologically discriminate h5-HT1B and h5-HT1D receptors. *Naunyn. Schmiedebergs Arch. Pharmacol.* **356**, 312–320 (1997).
34. Storer, R. I. *et al.* Multiparameter Optimization in CNS Drug Discovery: Design of Pyrimido[4,5- *d*]azepines as Potent 5-Hydroxytryptamine 2C (5-HT $_{2C}$) Receptor Agonists with Exquisite Functional Selectivity over 5-HT $_{2A}$ and 5-HT $_{2B}$ Receptors. *J. Med. Chem.* **57**, 5258–5269 (2014).
35. Wiltschko, A. B. *et al.* Revealing the structure of pharmacobehavioral space through motion sequencing. *Nat. Neurosci.* **23**, 1433–1443 (2020).
36. Kolda, T. G. & Bader, B. W. Tensor Decompositions and Applications. *SIAM Rev.* **51**, 455–500 (2009).
37. Williams, A. H. *et al.* Discovering Precise Temporal Patterns in Large-Scale Neural Recordings through Robust and Interpretable Time Warping. *Neuron* **105**, 246-259 e8 (2020).
38. Wang, Q. *et al.* The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell* **181**, 936-953 e20 (2020).
39. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
40. Ogawa, S. K., Cohen, J. Y., Hwang, D., Uchida, N. & Watabe-Uchida, M. Organization of Monosynaptic Inputs to the Serotonin and Dopamine Neuromodulatory Systems. *Cell Rep* **8**, 1105–1118 (2014).
41. Pollak Dorocic, I. *et al.* A whole-brain atlas of inputs to serotonergic neurons of the dorsal and median raphe nuclei. *Neuron* **83**, 663–78 (2014).
42. Ren, J. *et al.* Anatomically Defined and Functionally Distinct Dorsal Raphe Serotonin Sub-systems. *Cell* **175**, 472-487 e20 (2018).
43. Weissbourd, B. *et al.* Presynaptic partners of dorsal raphe serotonergic and GABAergic neurons. *Neuron* **83**, 645–62 (2014).
44. Ren, J. *et al.* Single-cell transcriptomes and whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. *Elife* **8**, e49424 (2019).
45. Swanson, L. W. Cerebral hemisphere regulation of motivated behavior. *Brain Res* **886**, 113–164 (2000).
46. Belmaker, R. H. & Agam, G. Major depressive disorder. *N Engl J Med* **358**, 55–68 (2008).
47. Ravindran, L. N. & Stein, M. B. The pharmacologic treatment of anxiety disorders: a review of progress. *J Clin Psychiatry* **71**, 839–54 (2010).

48. Ragan, T. *et al.* Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nat Methods* **9**, 255–8 (2012).
49. Renier, N. *et al.* Mapping of Brain Activity by Automated Volume Analysis of Immediate Early Genes. *Cell* **165**, 1789–802 (2016).
50. Kaltenecker, D. *et al.* Virtual reality-empowered deep-learning analysis of brain cells. *Nat. Methods* **21**, 1306–1315 (2024).
51. Achard, C. *et al.* CellSeg3D: self-supervised 3D cell segmentation for microscopy. Preprint at https://doi.org/10.1101/2024.05.17.594691 (2024).
52. Hansen, H. H. *et al.* Whole-brain activation signatures of weight-lowering drugs. *Mol. Metab.* **47**, 101171 (2021).
53. Davoudian, P. A., Shao, L.-X. & Kwan, A. C. Shared and Distinct Brain Regions Targeted for Immediate Early Gene Expression by Ketamine and Psilocybin. *ACS Chem. Neurosci.* **14**, 468–480 (2023).
54. Friedmann, D. *et al.* Mapping mesoscale axonal projections in the mouse brain using a 3D convolutional network. *Proc Natl Acad Sci U A* **117**, 11068–11075 (2020).
55. Batty, E. *et al.* BehaveNet: nonlinear embedding and Bayesian neural decoding of behavioral videos. *Adv. Neural Inf. Process. Syst.* **32**, (2019).
56. Pandarinath, C. *et al.* Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat. Methods* **15**, 805–815 (2018).
57. Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360–368 (2023).
58. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).