Published in final edited form as:

Nature. 2021 February ; 590(7847): 606–611. doi:10.1038/s41586-020-03061-2.

Evolving schema representations in orbitofrontal ensembles during learning

Jingfeng Zhou[1,*], **Chunying Jia**[2], **Marlian Montesinos-Cartagena**[1], **Matthew P.H. Gardner**[1], **Wenhui Zong**[1], **Geoffrey Schoenbaum**[1,*]

[1]Intramural Research Program of the National Institute on Drug Abuse, Baltimore MD, USA

[2]Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, USA

Abstract

How do we learn about what to learn about? Specifically, how do the neural elements in our brain generalize what has been learned in one situation to recognize the common structure of – and speed learning in – other, similar situations? We know this happens; we and our mammalian kith become better at solving new problems – learning and deploying schemas [1–5] – as we go through life. However, we have little insight into this process. Here we show that using prior knowledge to facilitate learning is accompanied by the evolution of a neural schema in the orbitofrontal cortex (OFC). Single units were recorded from rats deploying a schema to learn a succession of odor sequence problems. With learning, OFC ensembles converged on a low-dimensional neural manifold across both problems and subjects; this manifold represented the common structure of the problems and its evolution accelerated across their learning. These results demonstrate the formation and use of a schema in a prefrontal brain region to support a complex cognitive operation. Our results not only reveal an important new role for the OFC in learning but also have significant implications for using ensemble analyses to tap into complex cognitive functions.

Rats were trained on an odor sequence task [6,7]. On each trial, they sampled one of 16 odors and decided whether to respond to obtain reward (Fig. 1a). The odors were organized into two sequence pairs (Fig. 1b; S1a/b, S2a/b), each with 6 trials (positions; P1–6). All but two odors had a fixed association with reward, so for most of the 24 trial types information about sequence was not required, although it could be used to anticipate reward. However, for the odors at P4 and P5 in S2 (S2a4+, S2b4−, S2a5−, and S2b5+), the meaning of the same odor differed between the two subsequences (S2a and S2b). At these positions, the rats had to maintain the sequence information across several trials to perform correctly.

---

*Correspondence or requests for materials should be addressed to J.Z. (jingfeng.zhou@nih.gov) or G.S. (geoffrey.schoenbaum@nih.gov).

AUTHOR CONTRIBUTIONS

J.Z. and G.S. designed the experiments; J.Z. and M.M.-C. collected the data; J.Z. analyzed the data with advice and technical assistance from C.J., M.P.H.G., and W.Z.; and J.Z. and G.S. wrote the manuscript with input from the other authors.

COMPETING INTERESTS

The authors declare no competing interests.

After shaping, the rats were trained on five new problems. In each, the sequence and reward structure (Fig. 1b) remained the same, but 16 new odors were used (Problems #1 – #5; Fig. 1c). Thus, for each problem, the rats had to learn how the new odors fit into the sequence template (e.g. the schema) learned in prior problems. Rats trained 15 – 23 days on each problem. To align the analyses, each dataset was truncated to 15 sessions, including the first 14 sessions of learning and data from the session with the best performance thereafter (Extended Data Fig. 1). Behaviors reflecting the sequences emerged during this training, including correct responding on trial types that required recall of the sequences (Fig. 1d–g, Extended Data Fig. 2) as well as changes in the time the rats spent to initiate each trial – their poke latency – reflecting reward availability on the current and upcoming trials (Fig. 1h–j).

Reduced neural dimensionality

The numbers of recorded neurons (n = 1122.9 ± 41 neurons on each day, all problems and rats combined; Extended Data Fig. 3a–b, Supplementary Tables 1–2) and their overall activity remained stable across days (Extended Data Fig. 3c–d). The percent of neurons that were selective among the 24 trial types decreased during learning, while the percent of neurons that were sensitive to reward increased (Extended Data Fig. 3e–f). We used an independent component analysis (ICA) followed by a linear discriminant analysis (LDA) to reduce the dimensions in the single unit data (480 trials × N; N = ~1123 neurons per day × 4 epochs × 8 time points; 4 epochs: “light”, ”poke”, ”odor”, ”unpoke”; peri-event time at each epoch: −0.2 – 0.6 s; bin = 0.1 s; Fig. 2) and then measured the Mahalanobis distance (i.e., representational dissimilarity) between each pair of trial types within the resultant activity space (i.e., linear discriminant components; LCs). This revealed a drastic compression of the activity space from Day 1 to 15 (Fig. 2a–b, Extended Data Fig. 4a–h). The distribution of variance on the LCs shifted across days of training from a high dimensional pattern to one associated with lower dimensionality (Fig. 2c–d, Extended Data Fig. 4i–j). As a result, during learning, the percent of variance explained by the first three LCs increased (Fig. 2e); and the number of LCs that explained 80% variance decreased (Fig. 2f), especially for the S2 sequence, where information had to be remembered for accurate performance (Fig. 2g). These changes are consistent with the formation of a neural schema.

Evolving cross-problem/subject decoding

The observed neural representation on each problem is a projection of the true, underlying or “latent” neural representation within a particular subspace defined by currently recorded neurons [8]. If the activity in OFC is converging onto a schema with learning across problems, then this latent neural manifold (i.e., the true neural representation of the common structure of the task) should become increasingly similar with learning across each problem. However, changes in the recorded neurons across days and problems can cause the observed neural representations to appear misaligned, even if they come from the same latent neural manifold. To overcome this problem, we used a canonical component analysis (CCA) to align the neural activities acquired on different days and problems (i.e., manifold alignment [8]). The CCA transforms the dimensionality-reduced neural data from two odor problems such that they are maximally correlated across a set of dimensions or canonical components

(CCs). This resulted in two sets of paired CCs, describing the neural activities on the two odor problems, respectively. A higher correlation between a given pair of CCs means better alignment of the two datasets on this particular dimension, consistent with a more generalized, as opposed to idiosyncratic, neural representation across the two odor problems (Extended Data Fig. 5a).

If a neural schema exists, this procedure should reveal alignment of the neural activity across problems; the stronger the schema, the stronger the alignment across problems; further we should be able to use the CCs derived from this analysis to identify or map the trial types from the CCs derived from the analysis of other problems. To test these predictions, we conducted a decoding analysis in which we used 60 CCs generated from two of the five problems as training sets, and 60 CCs generated from two of the remaining problems as test sets (Extended Data Fig. 5b). Comparing the corresponding CCs, we found strong correlations between the CCs derived from the training and test data, with a notable increase from Day 1 to Day 15, consistent with the development of a generalized neural code – the neural signature of a schema – across problems (Fig. 3a, Extended Data Fig. 6).

A closer examination of the activity of the paired CCs revealed that the first three exhibited a clear relationship to specific task features: “Current Value” defined by whether the trial was rewarded, “Odor Overlap” defined by whether the odor was unique or shared, and “Positional Alternation” defined by an alternating pattern along each sequence. The relationships between the first three CCs and their respective task features increased in strength from Day 1 to Day 15 (Figs. 3a–c), generally improving cross-problem decoding of the trial types (Fig. 3d–e; aligned). Importantly, the accuracy of this cross-problem decoding was disrupted when the order of the 24 trial types was shuffled independently in each problem to misalign the sequence structure (Fig. 3d–e; misaligned); this demonstrates the dependence of the decoding on the generalizable structure of the problems, direct evidence for the construction of a neural representation of the odor sequence task capable of generalizing across problems. This process of construction was also evident in the marked changes in the cross-problem decoding that occurred with learning (Fig. 3d; aligned), particularly after the first few days of training (Fig. 3f), which reflected bidirectional changes in the ability of the decoder to correctly distinguish sequences (Extended Data Fig. 7a–b) and positions (Extended Data Fig. 7c–f). Specifically, cross-problem decoding of the positions improved during learning, whereas cross-problem decoding within each position tended to diminish for odor-unique positions (behaviorally-similar but discriminable by odors; P1–2), while improving at odor-overlapping positions (behaviorally-distinct but discriminable only based upon previously presented odors; P4–6, S2).

While the changes in cross-problem decoding presumably reflect convergence of the neural representations across problems within each subject, it is not clear from this analysis whether neural representations in the OFC also converged on a common solution or schema across subjects. Though not necessary, the ability to decode across subjects would imply the presence of a common neural representation or schema. To address this question, the rats were divided into groups and rats in one were used to generate CCs for training sets, while rats the other were used to generated CCs for test sets (Extended Data Fig. 5b). Again, we found strong correlations between CCs from the training and test data (Fig. 4a, Extended

Data Fig. 8); again the first three CCs developed increasingly strong relationships to the same three task features (“Current Value”, “Odor Overlap”, and ”Positional Alternation”) during training (Fig. 4a–c); and again these and the other CCs allowed successful decoding of the different trial types (Fig. 4d–e), which evolved with training (Fig. 4f, Extended Data Fig. 9).

Accelerating learning and neural changes

Poke latency measures how quickly the rats initiate a trial after the house-lights come on; rats responded more quickly when they expected more reward in the future than when they expected less reward (Fig. 1h–j). This pattern developed with learning on each problem, and the speed at which the pattern emerged during learning became faster across problems (Fig. 5, Extended Data Fig. 10). This accelerated learning is evidence of the operation of a schema reflecting knowledge of the sequences. While the temporal resolution of our neural data analysis demonstrating schema formation is, by necessity, different from resolution of the behavioral analysis demonstrating schema operation, the two should be related if the schema is manifest in the neural activity in OFC. Consistent with this idea, we found that both dimensionality reduction (Fig. 5c) and schema evolution (Fig. 5d) accelerated similarly.

The importance of identifying schemas

Learning what to learn about and generalizing from one situation to another is arguably one of the most fundamental abilities that distinguishes higher intelligence. The effects of the resultant schemas can be seen in simple motor or sensory processing, where learning one skill facilitates acquisition of another – knowing how to ski helps us learn to snowboard – but they are also evident in experience-based improvements in more abstract problem-solving skills. From buying groceries to planning experiments, we get better at navigating new situations if we have experienced similar ones in the past.

Here we identified a neural signature of schema formation in rats using sequences of odors to effectively retrieve rewards. As they were given new sets of odors, the rats became faster to exhibit behavior reflecting a knowledge of the sequences. Against this backdrop, we found that neural activity in OFC converged on a common manifold or representation of the task. This convergence was evident in a reduction in the number of dimensions characterizing the manifold and in improved manifold alignment across both problems and subjects, neural changes that accelerated with accelerated learning. Importantly, although these neural changes were correlative and thus likely to track increases in efficiency of motor behaviors and other covariates [9], the rats were pre-trained on the basic skills required in the task, and the improved alignment in neural activity reflected improved representation of critical task-relevant cognitive concepts, some of which were idiosyncratic features of the odor sequence problems. This work joins a small but growing number of other studies that have also found neural correlates of schemas in other brain regions [10–16], including generalization of single unit and ensemble firing patterns to incorporate new exemplars of a class [11].

The identification of a neural signature of schema formation in the OFC is intriguing in light of this area's involvement in flexible problem solving, learning set formation, and even generalization [17–19]. A common thread across these settings is the need to use the causal structure of the relationships defining the task at hand to support normal behavior [20]. This observation has led to the hypothesis that the OFC is part of a circuit, including other prefrontal areas and hippocampus, required for cognitive mapping [6,20–24]. The current correlative data suggest that the OFC may play a specialized role within that circuit, allowing the meta-structure or rules that characterize individual maps to be abstracted and deployed to aid the formation of new maps. This more specialized function would be consistent with well-appreciated data showing that the OFC is necessary when behavior requires a cognitive map and with less well-appreciated evidence that this is only true when *new* information – a new map – is necessary [25]. This is evident in devaluation [26], sensory preconditioning [18], and over-expectation [27] tasks, where the OFC is required in phases where an established map must be adjusted to reflect new information. By contrast, when behavior does not require the integration of new information, in these as well as other settings [28–30], the OFC tends not to be necessary.

More broadly, our current report identifies a neural signature of schema formation in a prefrontal brain region and for information that is not easily characterized as simply reflecting optimization of motor or sensory processing. The pattern of neural activity in the OFC, identified using dimensionality reduction and manifold alignment, converged onto a common solution for a complex cognitive operation. Further, the resultant cognitive schema was neurally-generalizable not only across problems within the same brain, but also across brains [14]. That the OFC networks in different subjects converged on a common manifold is significant, since it implies that individuals organize information in a similar format, even for advanced prefrontal functions. This is important for neuroscience, since it suggests we can identify even these types of functions at a granular level by studying groups of subjects. It also has implications for utilizing brain-machine interfaces to enhance functions beyond sensory perception or simple motor activity, since if neural schemas for complex cognitive operations are not idiosyncratic to individuals, then they may be defined from a general understanding or knowledge of their shape, rather than needing to be tailored to each individual on each problem, which would be much less practical.

METHODS

Rats

Subjects were nine male Long-Evans rats (Charles River, 175 – 200 g, ~3-month-old) housed individually on a 12-h light/dark cycle and given *ad libitum* access to food in an animal facility at the AAALAC-accredited animal care facility at the National Institute on Drug Abuse Intramural Research Program (NIDA-IRP). Water was removed the day before any testing day, and they received free access to water for 10 min in their home cages each afternoon after testing. If there was no testing the next day, the rats were given free access to water. All behavioral testing was carried out at the NIDA-IRP. Animal care and experimental procedures complied with the U.S. National Institutes of Health (NIH) guidelines and were approved by the Animal Care and Use Committee (ACUC) at the NIDA-IRP.

Odor sequence task

The training was conducted in aluminum boxes (~18” on a side) equipped with a port for odor delivery and a well for delivery of sucrose solution. Task events were controlled by a custom-written C++ program and a system of relays and solenoid valves; entries into the odor port and the fluid well were detected by infrared beam sensors. The availability of each trial was signaled by the illumination of two house-lights located on the wall above the odor port. Nosepoke into the odor port within 5 seconds after light onset initiated the trial, leading to odor delivery after a 500-ms delay. Rats were required to remain in the port for an additional 500-ms; the trial was aborted, and the lights extinguished if the rat left the odor port before this time had elapsed. After 500-ms, the rats were free to leave the port, which terminated odor delivery. After port exit, rats had 2-s to respond at the fluid well. On rewarded trials, a response led to the delivery of a sucrose solution (10% w/v; 50 μL) after a random delay ranging from 400 to 1500-ms. Upon exit from the well, non-responding during the 2-s period, or responding on non-rewarded trials, the house lights were extinguished, indicating the end of the trial and the beginning of the inter-trial interval (ITI). A 4-s ITI followed correct trials, and an 8-s ITI followed trials on which the rat made an error.

On each trial, one of 16 odors was delivered to the odor port. The 16 odors were organized into two pairs of sequences (S1a, S1b, S2a, and S2b), described below. The odor identity is indicated by a number, and reward and non-reward is indicated by the positive (+) and negative (−) symbols, respectively:

$$\text{S1a: } 0+\ 1-\ 4-\ 5+\ 6-\ 7+$$

$$\text{S1a: } 2+\ 3-\ 4-\ 5+\ 6-\ 7+$$

$$\text{S2a: } 8+\ 9-\ 12-\ 13+\ 14-\ 15+$$

$$\text{S2b: } 10+\ 11-\ 12-\ 13-\ 14+\ 15+$$

Before training with any odors, rats were first shaped to nosepoke at the odor port and then respond at the well for a reward. Then, they were trained to discriminate a single odor pair (one rewarded and one non-rewarded odor) from sequence S1a or S1b. Sessions consisted of a maximum of 480 trials. After rats reached a criterion of > 90% correct performance on this initial odor pair, additional odor pairs were added until the rats were able to perform accurately (> 75% correct) on every trial type in a session containing all of the odors in sequences S1a and S1b. After meeting this criterion, rats were then trained on odor pair 13/14 from sequence S2, including several reversals. After the third reversal, additional odor pairs were added from sequence S2 if the rats were able to perform accurately (> 75% correct) on every trial type in a session containing all of the odors in sequences S2a and S2b. Once sequence S2 had been fully introduced in this manner, the rats began sessions

containing the full set of sequences (S1a/b and S2a/b). In this final training phase, each sequence was repeated 20 times to make 480 trials in total. Sequences S1a and S1b were always followed by S2a or S2b with roughly equal probability (0.55 and 0.45, respectively). Sequence S2a and S2b were always followed by S1a or S1b also with slightly more dissymmetry in probability (0.67 and 0.37, respectively). The overall sequence was repeated from start to finish in each session.

S1b S2a S1a S2a S1a S2b S1b S2b S1b S2b S1b S2a S1b S2b S1a S2a S1a S2b S1a S2a S1a S2b S1b S2a S1b S2a S1b S2b S1b S2a S1b S2b S1a S2a S1a S2b S1a S2a S1a S2b S1a S2a S1b S2a S1b S2a S1a S2b S1a S2a S1a S2a S1a S2b S1b S2b S1b S2b S1b S2b S1b S2a S1a S2a S1a S2b S1b S2b S1b S2b S1b S2a S1b S2b S1a S2a S1a S2b S1a S2a

Rats trained on the full set of sequences until they were able to perform accurately (> 75% correct) on every trial type in a session, then electrode arrays were implanted bilaterally in the OFC.

Surgical procedures

Rats were implanted with two drivable bundles of 16 electrodes (32 electrodes in total), made from nickel-chromium wires (25 μm in bare diameter; AM Systems, WA), in bilateral OFCs (AP: 3 mm, ML: 3.2 mm). Each wire bundle was housed in a stainless-steel hypodermic tubing (27 Gauge, 0.01625” OD, 0.01025” ID) and cut with a pair of fine bone-cutting spring scissors (#16144–13; Fine Science Tools; Foster City, CA) to extend 1.5 – 2 mm beyond the end of the tubing inside the brain. The tips of wires were initially placed at 4.2 mm ventral from the brain surface. After surgery, rats were given Cephalexin (15 mg/kg) orally twice a day for two weeks to prevent any infection.

New odor problems

After the pre-training with the first odor set, electrode implantation, two weeks of recovery, and one-week of reminder training on the original odor problem, the rats were successively given five new odor problems (#1 – #5) to learn. These problems consisted of 16 new odors, arranged in sequences with the same structure as the original odor problem. Single-unit activity was recorded bilaterally from OFC during learning of each new problem, which took from 15 – 23 days. The neural dataset was truncated to 15 days on each problem. Data on Day 1 – 14 were from the first 14 days of training on each problem; data on Day 15 was taken from the day of the best performance between Day 15 – 23.

Single-unit recording

Electrophysiological signals were recorded with the Plexon Multichannel Acquisition Processor (MAP) systems (v2.7.0; Plexon, Dallas, TX). The initial signals collected by the electrodes were sequentially amplified through a headstage (20×), a differential preamplifier (50×), and a final acquisition processor (1 – 32×). A band-pass filter (250 – 8, 000 Hz) was used to isolate spike activities, and a threshold was set manually for each active channel to capture unsorted spikes. Timestamps for behavioral events were sent to the Plexon system, synchronized and recorded alongside the neural activity.

Spikes were sorted later offline to remove noise and isolate single units using Offline Sorter (v4.0; Plexon, Dallas, TX) with a built-in template matching algorithm. In short, in a 2-D view of spike features (e.g., PC1, PC2, PC3, spike valley, non-linear energy), a small group of spike waveforms were manually selected to generate an averaged spike template. In each channel, one or more spike templates were created based on their waveform shapes, visual discriminability, and their signal-to-noise ratios. Other waveforms were then assigned to clusters based on their distances (i.e., sum-of-squares difference) to all the templates with an adjustable fit tolerance. The waveform was assigned to one cluster if its minimal distance was less than the fit tolerance for the template. Sorted files were saved as NeuroExplorer (v4.0; Nex Technologies, Colorado Springs, CO) formatted files to extract unit and behavioral event timestamps, which were then exported to MATLAB (2019a; MathWorks, Natick, MA) for further analysis.

Between odor problems, the electrodes were advanced ~120 μm in order to change the neural population being sampled. Within a given problem, the electrodes were not advanced. However, we make no claims regarding whether single units recorded on different days within the same problem are the same or different neurons.

Quantification and statistical analyses

The number of rats and the number of neurons were not predetermined by any statistical methods but are comparable to those reported in previous publications from our and other labs [6,8,31–33]. All data were analyzed using MATLAB (2019a; MathWorks, Natick, MA).

Peri-event spike dynamics

Each trial was divided into 6 epochs associated with different task events: “light”, “poke”, “odor”, “unpoke”, “choice”, and “outcome”. On rewarded trials, the time of well-entry was labeled as “choice”. “Outcome” was at the time of reward delivery. On non-rewarded trials, the end of the 2-second window for responding was labeled as “choice” and a time point 1.5 s after the “choice” as “outcome”. Behavioral performance was quantified by the percent of trials on which the rats responded correctly (“%correct”), their reaction time from the odor port to the fluid well (“reaction time”), and the latency with which they initiated a trial after light onset (“poke latency”). The spike train for each isolated single unit was aligned to the onset of each task event for the calculation of a peri-event time histogram (PETH). Pre-event time was set to be 200 ms, and post-event time 600 ms. Spike number was counted with a bin = 100 ms. A Gaussian kernel (σ = 50 ms) was used to smooth the PETH on each trial. The PETH of each neuron on each task event consisted of 480 trials (rows) and 8 time points (columns).

For neural ensemble analyses, we concatenated PETHs of all neurons across four task epochs (“light”, “poke”, “odor”, “unpoke”), which resulted in a matrix, $X \in \mathbb{R}^{M \times T}$, Where M is the number of trials (480 trials), and T is the number of concatenated neurons across time (number of neurons × 4 epochs × 8 time points). For each repeat of neural ensemble analyses (n = 500 repeats), the trial order within each trial type (24 trial types) for each neuron was shuffled to generate a pseudo-ensemble, which also removes the temporal correlation between neurons within the same trial type.

Dimensionality reduction

To reduce the dimensionality of the neural data, we used an independent component analysis (ICA; MATLAB toolbox: GroupICATv4.0b) [34]. ICA is a blind source separation technique that attempts to decompose an observed multivariate signal into independent source components, which is commonly used as a dimensionality reduction tool to obtain a low-dimensional representation of high-dimensional original data [35–38]. ICA was implemented on a subspace resulted from a principal component analysis (PCA) to reduce the effect of noise.

For the neural data $X \in \mathbb{R}^{M \times T}$, where M is the number of trials (M = 480 trials), and T is the number of concatenated neurons, we first ran a PCA on the trial dimension to extract a signal subspace $Y \in \mathbb{R}^{N \times T}$, where N is the number of principal components (PCs) used for further analyses (N = 100). The eigenvector matrix is denoted as $V \in \mathbb{R}^{M \times N}$. The noiseless ICA model can be written as

$$Y = AS,$$

where $A \in \mathbb{R}^{N \times N}$ denotes the mixing matrix and $S \in \mathbb{R}^{N \times T}$ denotes the latent source matrix. We used the Infomax [34,39], a widely used ICA algorithm, to estimate a demixing matrix $W \in \mathbb{R}^{N \times N}$ such that the components of the estimated source matrix $\hat{S} \in \mathbb{R}^{N \times T}$ are statistically independent. $\hat{S}$ is computed as $\hat{S} = WY$. The estimated mixing matrix $\hat{A} \in \mathbb{R}^{N \times N}$ is the inverse of W: $\hat{A} = W^{-1}$. Back-reconstruction was performed on $\hat{A}$ with the previous PCA transformation to generate $B \in \mathbb{R}^{M \times N}$: $B = V^{+T}\hat{A}$, where $(\cdot)^{+T}$ denotes the transpose of the Moore-Penrose pseudo-inverse of a matrix. Each column of B, representing the weights of the corresponding estimated source $\hat{S}$, is referred to as the trial covariation. The trial covariation matrix B was used as a low-dimensional representation of the high-dimensional neural data.

Due to the variability of the ICA algorithms in the solution space, we ran the ICA on each dataset 20 times and selected the most consistent run by using an algorithmic consistency metric based on the cross-inter-symbol interference (cross-ISI) [40]. The ISI is a global metric for performance evaluation but only when the ground-truth is known. A smaller ISI between a single run and the ground-truth means the estimation is closer to the ground-truth. The metric cross-ISI has been developed to overcome the challenge when ground-truth is not available for real-world data. In brief, the current run is assumed as the ground-truth, the cross-ISI of the current run is computed as an average of all ISIs between the current run and other runs. The run with the lowest cross-ISI (i.e., most consistent run) was selected for further analysis.

Representational dissimilarity analysis

A linear discriminant analysis (LDA) was applied to the low-dimensional representation of the neural data, $B \in \mathbb{R}^{M \times N}$. The LDA finds the components in the underlying subspace that best separate classes. Suppose $x_k \in \mathbb{R}^{1 \times N}$, k = 1, 2, 3, … , M, is a vector of the neural data B on Trial k (M = 480 trials, N = 100), c is the number of classes (c = 24 trial types), μ is the

mean for each class, $\bar{x}$ is the overall mean. The with-class scatter matrix S_w and between-class scatter matrix S_B are defined by:

$$S_W = \sum_c \sum_{i \in c} (x_i - \mu_c)(x_i - \mu_c)^T,$$

$$S_B = \sum_c (\bar{x} - \mu_c)(\bar{x} - \mu_c)^T.$$

LDA tries to maximize the ratio of between-class variance and within-class variance of the linearly transformed data:

$$\widehat{W} = \text{argmax}\left[\left(W^T S_B W\right)/\left(W^T S_W W\right)\right],$$

where W is the argument and $\widehat{W} \in \mathbb{R}^{N \times N}$ the resulted coefficient matrix for the LDA transformation. The LDA components are computed as $B_{LDA} = B\widehat{W}$, where $B_{LDA} \in \mathbb{R}^{M \times N}$.

The Mahalanobis distance was used to measure the distance or dissimilarity between each pair of trial-type means in the LDA subspace with B_{LDA}. The analysis was repeated for 500 times and the averaged dissimilarity matrix was presented as heatmap in the main figure. The 24 trial types were plotted in a 2-D scatter plot through the classic multidimensional scaling (also known as principal coordinates analysis; MATLAB function: *cmdscale*) method. For clustering analysis, an agglomerative hierarchical cluster tree was generated with the unweighted average distance method. The clustering results were plotted as dendrograms.

Dimensionality comparison

To estimate task-related neural dimensionality, we examined the distribution of variance on the LDA components [15]. Lower neural dimensionality means fewer neural components should be required to represent a fixed amount of the variance and a fixed number of components should be able to represent a larger amount of the variance. To test for this, we calculated variance explained by each neural component, and asked two questions: 1) How much variance can be explained by the first three neural components? 2) How many neural components are needed to explain 80% variance? To make comparisons of dimensionality reduction between problems, we 1) randomly selected 140 neurons from each odor problem on each day for each repeated analysis; 2) normalized the variance explained by the first three neural components across days (i.e., the explained variance on each day was divided by the peak variance found across days to ensure a 0 – 1 scale.) and denoted the numbers as “Dimensional Reduction Index”. A greater number means lower dimensionality.

Manifold alignment

We used the canonical correlation analysis (CCA) to align neural activities recorded in different sessions in a lower-dimensional space. CCA uses second-order statistics (i.e., cross-covariance) to discover the relationships between two sets of multidimensional

variables, by finding two sets of respective linear transformations (i.e. canonical coefficients), such that the correlation between two projected variables (i.e., canonical variables) is maximized. CCA has been used to align neural activities previously [8,41,42].

Consider $B^{[1]} \in \mathbb{R}^{M\times N}$ and $B^{[2]} \in \mathbb{R}^{M\times N}$ are the low-dimensional representations of the original neural data from two problems after the ICA (M = 480 trials; N = 100). To avoid overfitting of CCA, we performed PCA on $B^{[1]}$ and $B^{[2]}$ along the neuron dimension, which resulted in $C^{[1]} \in \mathbb{R}^{M\times L}$ and $C^{[2]} \in \mathbb{R}^{M\times L}$, respectively, where L is the number of PCs (L = 30) retained. CCA tries to find pairs of canonical coefficient vectors $W^{[1]} \in \mathbb{R}^{L\times L}$ and $W^{[2]} \in \mathbb{R}^{L\times L}$, such that the pairwise correlations between $U^{[1]} = C^{[1]}W^{[1]}$ and $U^{[2]} = C^{[2]}W^{[2]}$ are maximized. Columns in $U^{[1]} \in \mathbb{R}^{M\times L}$ and $U^{[2]} \in \mathbb{R}^{M\times L}$, ordered by paired correlation coefficients, are canonical variables or canonical components (CCs), which were used to represent aligned neural activities between problems.

Classification analyses

We used the LDA for multiclass classification and support vector machine (SVM; MATLAB toolboxes: libsvm-3.22 and ndt.1.0.4) [43,44] for binary classification. In general, classification accuracy was assessed by a leave-one-out cross-validation procedure. Specifically, on each repeat, one trial from each trial type was left out for future testing, and all the other trials were used to create the classifier. For repeats, the trial order for each neuron was shuffled within the same trial type. The trial-order shuffling was repeated for 500 times. For each time of trial-order shuffling, the leave-one-out cross-validation was repeated 200 times to estimate the decoding accuracy. The mean decoding accuracy for each trial type as shown in the confusion matrix was the mean across 500 runs (corresponding to 500 times of trial-order shuffling). The statistical significance of the mean decoding accuracy was determined by the 95% confidence interval estimated by running the same decoding process with label-shuffled data.

For each repeat in the cross-problem classification, the five odor problems were randomly assigned with numbers #1 – #5 (these numbers do not indicate the actual problem order). The training set came from one pair of odor problems (#1 and #2), and the test set came from another pair (#3 and #4). The neural data from the odor problem #5 was randomly combined with odor problem #1 or #2. For the training set, the original neural firing data from the two odor problems (#1 and #2) were separately concatenated along the neuron dimension and subjected to dimensionality reduction (ICA + PCA) as described above, which resulted in two matrices $B^{[1]} \in \mathbb{R}^{M\times L}$ and $B^{[2]} \in \mathbb{R}^{M\times L}$ (M = 480 trials, L = 30 PCs). The manifold alignment was performed on the two matrices to obtain aligned neural activities $U^{[1]} \in \mathbb{R}^{M\times L}$ and $U^{[2]} \in \mathbb{R}^{M\times L}$. $U^{[1]}$ and $U^{[2]}$ are concatenated along the neural dimension as the training data. With the same approach, $U^{[3]} \in \mathbb{R}^{M\times L}$ and $U^{[4]} \in \mathbb{R}^{M\times L}$ are the results of dimensionality reduction and manifold alignment performed on another pair of odor problems, #3 and #4. $U^{[3]}$ and $U^{[4]}$ were concatenated as the test data. If the correlation coefficient between two paired canonical components in the training and tests was negative, the sign of the component in the test set would be flipped. The cross-rat classification followed the same procedure except that the rats, instead of the odor problems, were divided into four groups (#1, #2, #3 and #4). Groups #1 and #2 were used to generate the training

set, and groups #3 and #4 were used to generate the test set. For cross-rat classification within each problem, 50 neurons were randomly selected from each rat with replacement for each repeat on each day.

For the correlation analyses between the CCA-aligned neural activities and three task features ("Current Value", "Odor Overlap", and "Alternating Position"), the dataset was not split into training and test sets. Specifically, for each repeat, the 5 problems (or 9 rats) were randomly divided into two groups (#1 and #2). We applied the same dimensional reeducation and manifold alignment process described above on the split datasets #1 and #2, which resulted in aligned neural activities: $U^{[1]} \in \mathbb{R}^{M \times L}$ and $U^{[2]} \in \mathbb{R}^{M \times L}$. The Pearson correlations between the first 13 canonical components in $U^{[1]}$ and the three task features were calculated.

Control analysis for cross-problem and cross-subject decoding

The control analysis was intended to mimic a control experiment, in which the components (number of odors, number of trials, likelihood of reward, etc.) remained the same across odor problems but the task structure – the sequence of odors and rewards – differed. In such a control experiment, odors would be presented randomly, with or without reward as in the current experiment, but without any sequence structure. Under these conditions, the trial types within each problem should remain decodable, but cross-problem and cross-subject classifications should fail. To approximate this condition with the current data, we shuffled the order of trial types within each problem before the manifold alignment, which effectively, although not fully, disrupted the sequence structure while leaving other components intact, and then we performed the above analyses to decode the trial types across problems and across subjects.

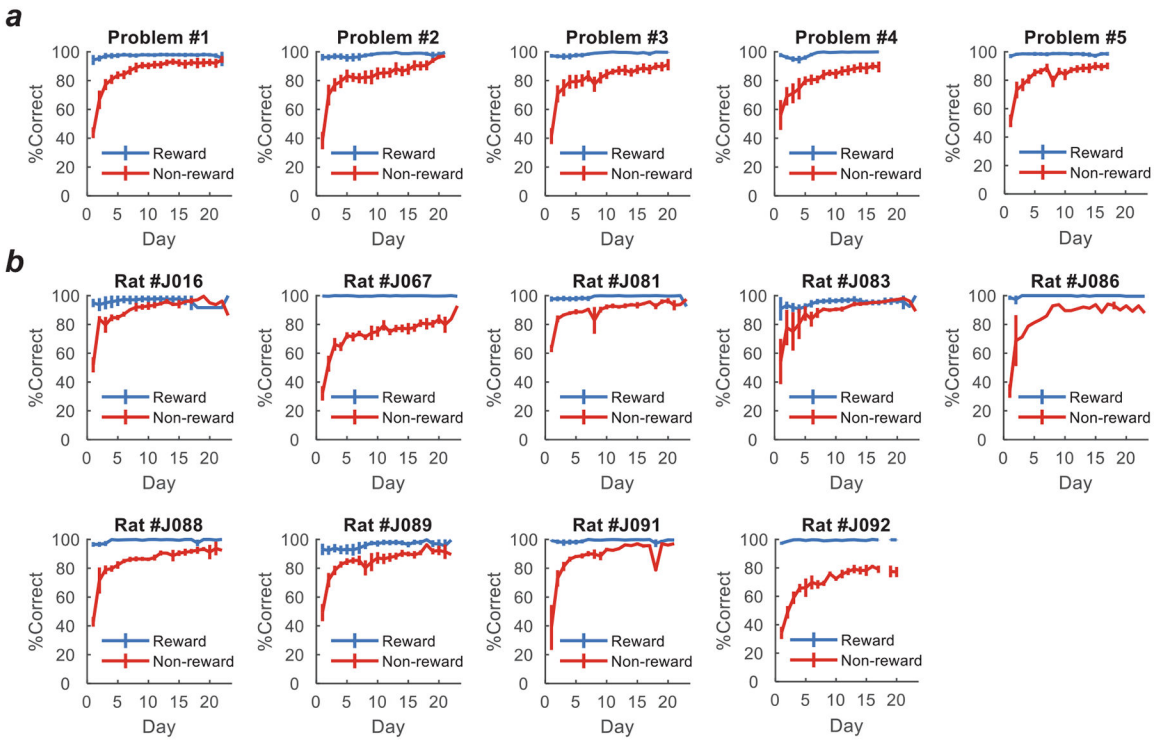
Data availability

The dataset used in this study is available at https://doi.org/10.17605/OSF.IO/5MH4Y

Code availability

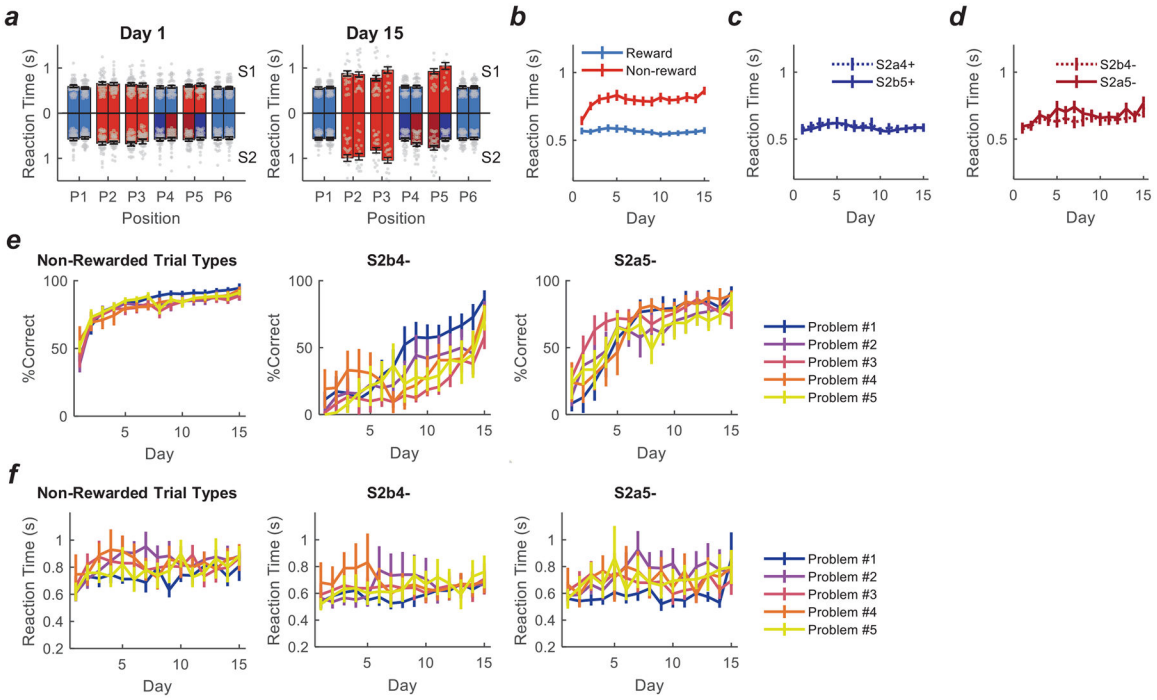
The MATLAB code used in this study is available at https://doi.org/10.17605/OSF.IO/5MH4Y

Extended Data



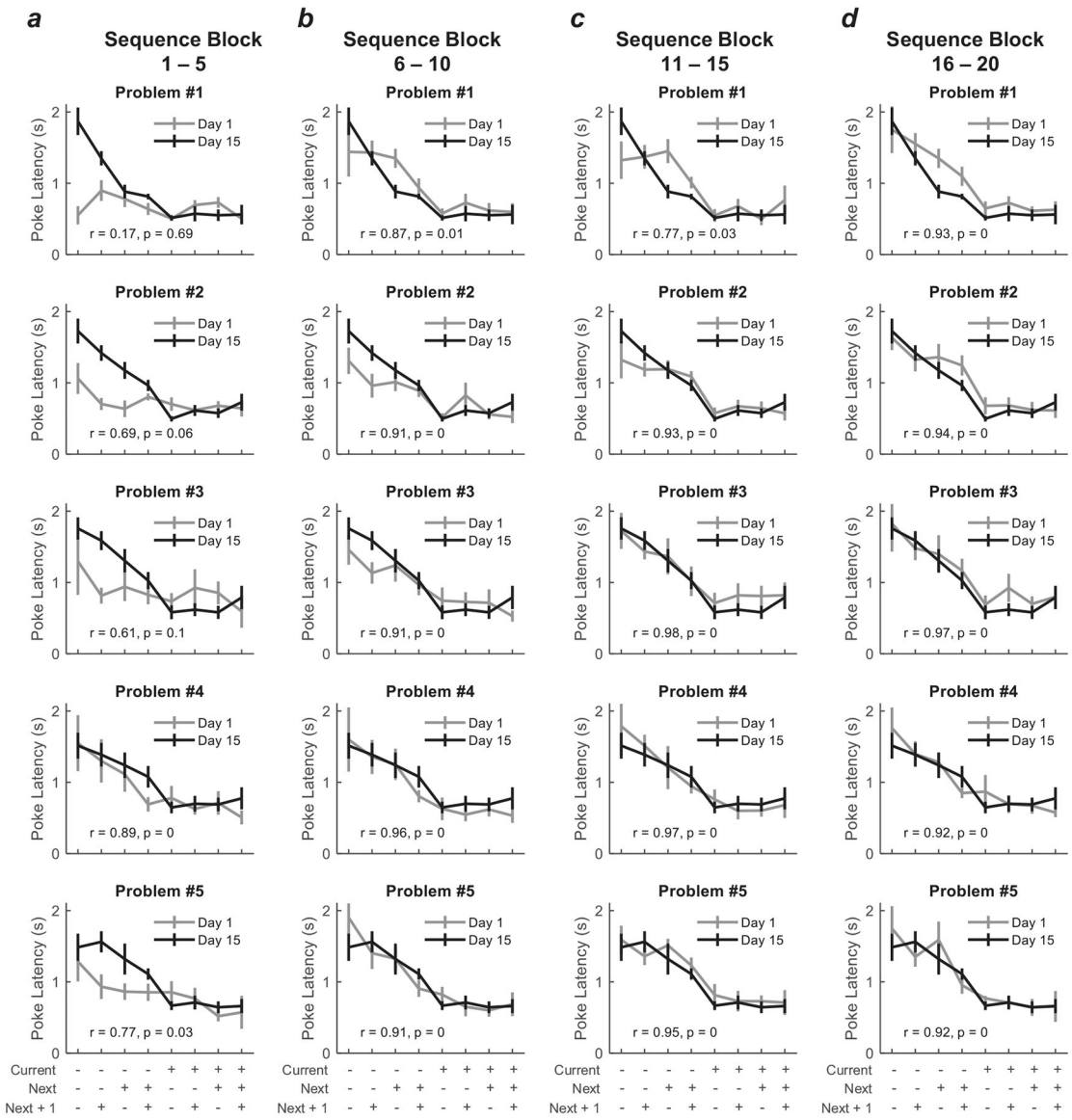
Extended Data Fig. 1 |. Behavioral performance on each problem and of each rat.

a, **b**, The behavioral learning was assessed by percent of correct (%correct) on rewarded trials ("Go" trials; blue) and non-rewarded trials ("No-Go" trials; red) across training days. Each rat accomplished one session each day. The data was plotted for each problem across rats (**a**; n = 9 rats) or each rat across problems (**b**; n = 5 odor problems). The days shown on the x-axis are the actual training days. Rats showed stable behavioral performance after Day 15 and not all rats finished 23 days of training. To align the learning process between problems for further data analyses, we truncated the learning on each odor problem to 15 sessions, consisting of data from the first 14 sessions of learning plus data from the session with the best performance thereafter. Note that the training day of the last sessions with the best performance was referred to as "Day 15", except in this figure. Data are presented as mean ± s.e.m. A two-way ANOVA was performed for each panel (Reward × Day; R × D). See Supplementary Table 8 for detailed statistics.



Extended Data Fig. 2 |. Reaction time and %correct during learning.

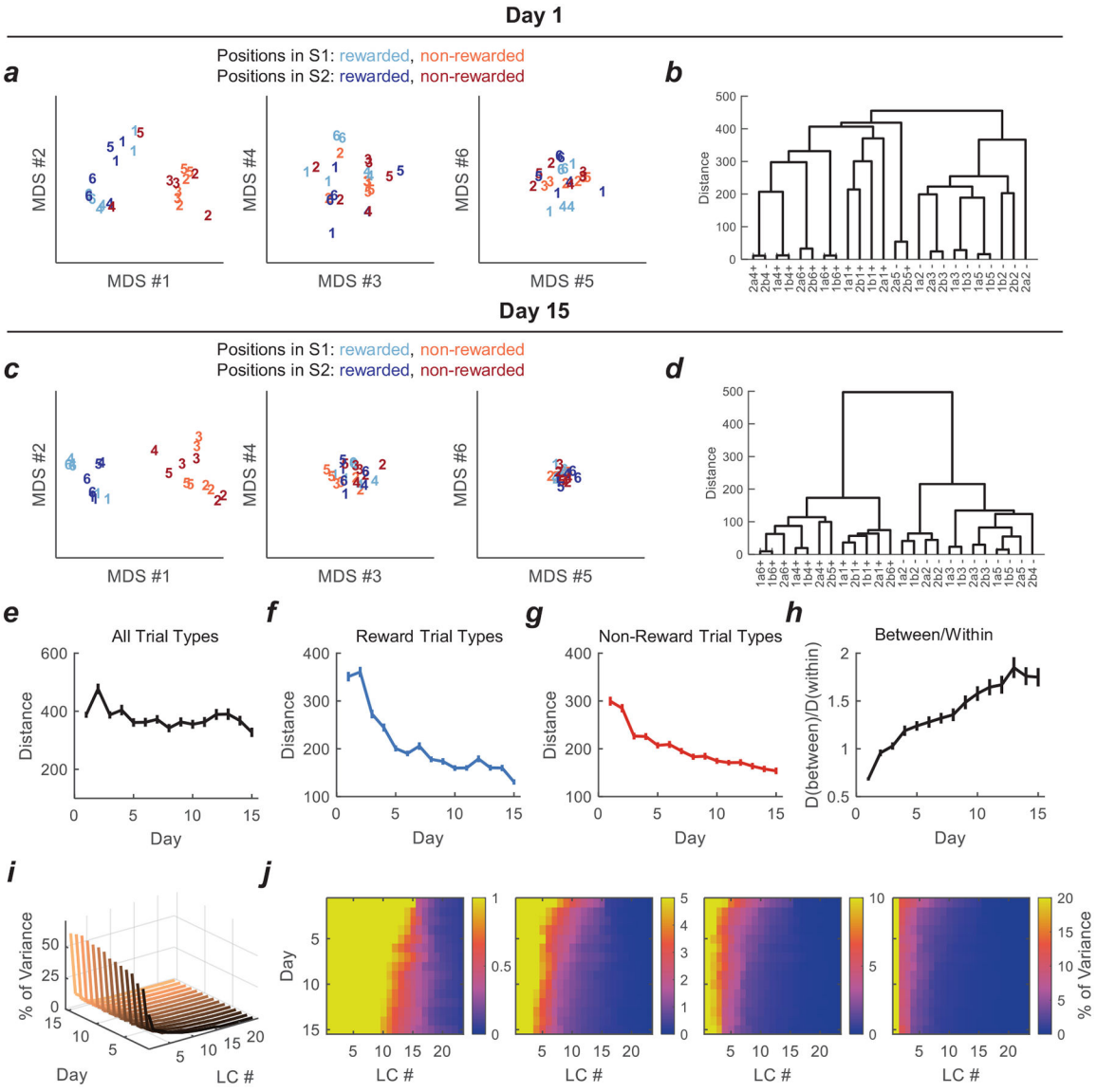
a, Reaction time on Day 1 and Day 15. The reaction time measured the time period from odor port exit (“unpoke”) to water well entry (“choice”). The data presented here only included correct rewarded trials and incorrect non-rewarded trials. Reaction times on trial types in sequence S1 are plotted upwards, and those on trial types in sequence S2 are plotted downwards. Darker colors highlight four trial types that require rats to remember and use past sequences to perform correctly. n = 37 and 36 sessions on Day 1 and Day 15, respectively. Data are presented as mean ± s.e.m. **b**, The changes of reaction time on correct rewarded trials and incorrect non-rewarded trials during learning. **c**, **d**, The changes of reaction time during learning on the two pairs of trial types (**c**: S2a4+ and S2b5+; **d**: S2b4− and S2a5−). **e**, %Correct on all non-rewarded trial types and two (S2b4− and S2a5−) that required the recall of odor sequences. **f**, Reaction time on all non-rewarded trial types and two (S2b4− and S2a5−) that required the recall of odor sequences. Only incorrect trials (rats making a “Go” choice on non-rewarded trials) were included. **b**-**f**, n = 37, 40, 40, 38, 38, 39, 38, 39, 39, 38, 39, 40, 36, 38, 36 sessions from Day 1 to Day 15. Data are presented as mean ± s.e.m. **a**-**d**, Two-way ANOVAs (Trial Type × Day). **e**-**f**, Two-way ANOVAs (Problem × Day). See Supplementary Table 9 for detailed statistics.



Extended Data Fig. 3 |. Histology and ingle-unit analyses.

a, Red squares show the reconstructed recording sites (n = 18 recording sites from 9 rats). Two electrode bundles were implanted bilaterally in OFC of each rat. Each electrode bundle consisted of 16 single wires. **b**, The number of neurons recorded across days. **c**, Cumulative distribution of neurons that showed different firing rates to all the odor stimuli. **d**, The averaged firing rates of all the neurons to all the odor stimuli. One-way ANOVA with the factor of Day: $F(14,16828) = 0.39$, $p = 0.98$. See **b** and Supplementary Tables 1 and 2 for n = number of neurons on each day. Data are presented as mean ± s.e.m. **e**, The percent of neurons that were significantly selective to at least one of the 24 trial types (one-way ANOVA; p values were adjusted by the Benjamini-Hochberg procedure to control the false discovery rate; BH-FDR; $p < 0.05$ was used to determine if one neuron was significantly selective to 24 trial types). **f**, The percent of neurons that showed selectivity to reward vs.

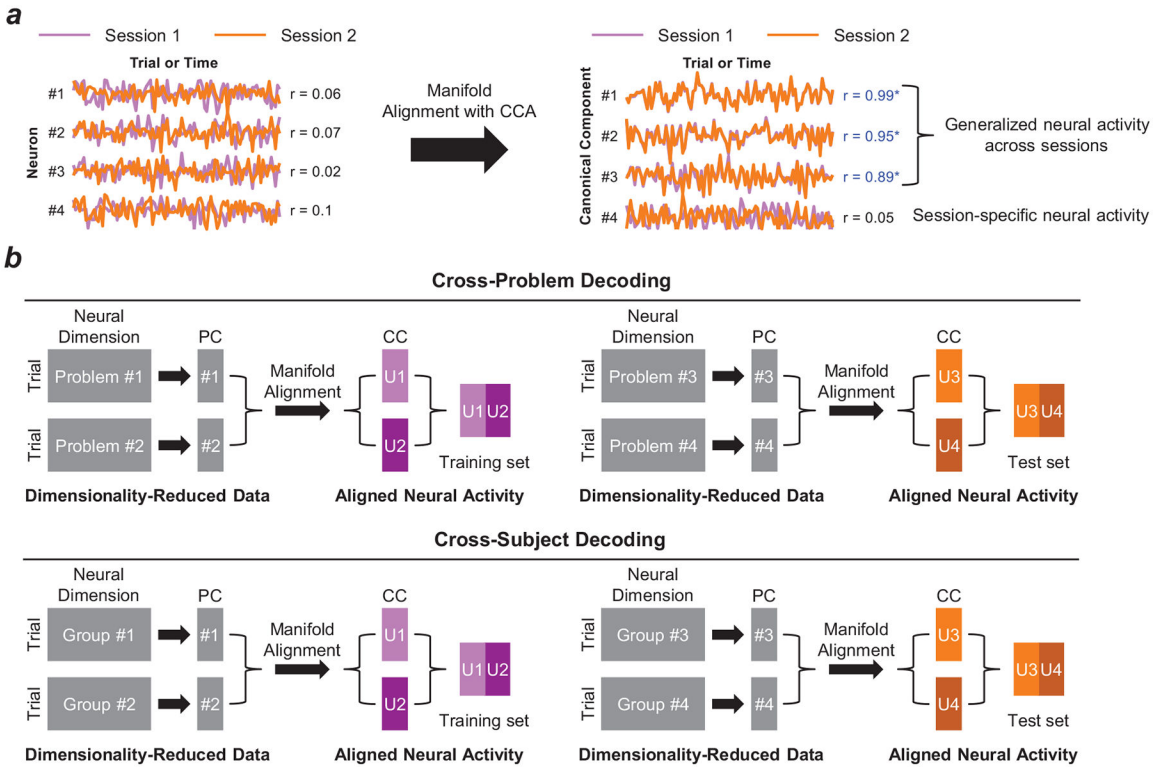
non-reward trials (one-way ANOVA; BH-FDR; $p < 0.05$ was used to determine if one neuron was significantly selective to the current value).



Extended Data Fig. 4 |. Changes of activity space during learning.

a, Using the classical multidimensional scaling (cMDS) to visualize the dissimilarity matrix shown in Fig.2. The colored numbers (1 – 6) indicate positions (P1 – P6). Light blue and light red: rewarded and non-rewarded positions in S1, respectively; dark blue and dark red: rewarded and non-rewarded positions in S2, respectively. **b**, Hierarchical clustering of 24 trial types based on population neural activities on Day 1. The dissimilarity matrix in Fig. 2 was used to construct a hierarchical clustering tree by an unweighted average linkage method. The clustering results were shown in dendrograms. **c**, The MDS plots to visualize the dissimilarity matrix on Day 15 with the same color code as in **a**. **d**. The hierarchical clustering analysis on Day 15. **e**, Averaged pair-wise distances between 24 trial types. One-way ANOVA (factor: Day): $F(14,7485) = 2279.2$, $p = 0$, $n = 500$ repeats. **f**, Averaged pair-

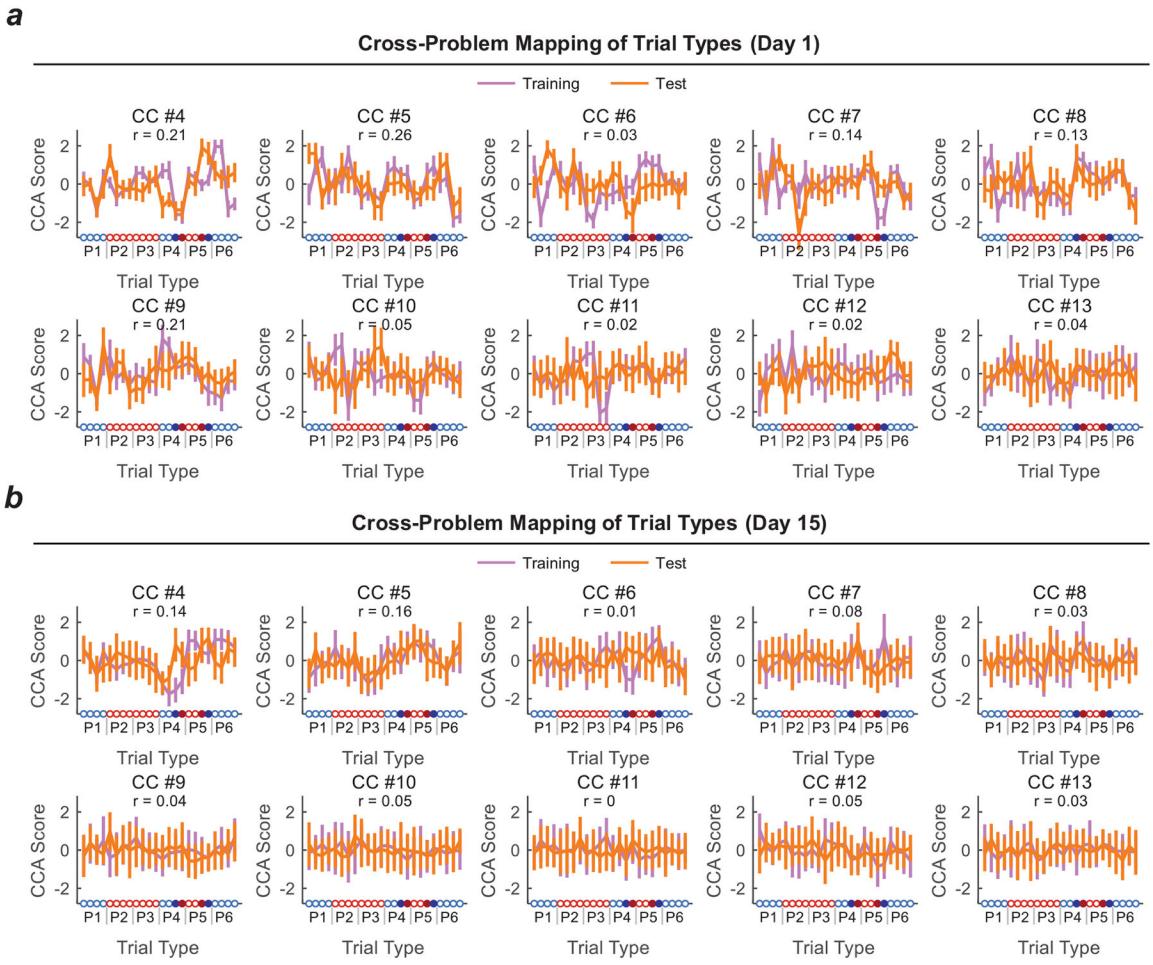
wise distances within reward trial types. One-way ANOVA (factor: Day): F(14,7485) = 4.4×10^{-4}, p = 0, n = 500 repeats. **g**, Averaged pair-wise distances within non-reward trial types. One-way ANOVA (factor: Day): F(14,7485) = 1.9×10^{-4}, p = 0, n = 500 repeats. **h**, The ratio of averaged pair-wise distance between and within reward vs. non-reward trials. One-way ANOVA (factor: Day): F(14,7485) = 1.1×10^{-4}, p = 0, n = 500 repeats. **e-h**, Data are presented as mean ± s.d. **i**, Percent of explained variance across linear discriminant components (LCs). **j**, Heatmap plots of variance distributions across LCs and training days. Warmer color mean higher percent of variance explained by certain LC, and *vice versa*. Four panels were used with different colored bar scales (from left to right: 0 – 1; 0 – 5; 0 – 10; 0 – 20) to better visualize the same result.



Extended Data Fig. 5 |. Obtaining training and test sets.

a, Simulation of manifold alignment with the canonical component analysis (CCA). Two sets of Gaussian signals (n = 4 for each set) were generated to represent two sets of neurons recorded from two respective task sessions. Correlations of paired neurons between sessions were controlled (r = 0.99, 0.9, 0.85, 0.01; Pearson correlation). The neurons between sessions were misaligned such that the correlations between paired neurons were low (shown in the left; p = 0.58, 0.52, 0.86, 0.35; Pearson correlation) to mimic the misalignment of neurons during experimental recording sessions. After the manifold alignment, the recovered neural components (i.e., canonical components) were aligned (p = 1.8×10^{-94}, 7.0×10^{-34}, 3.5×10^{-27}, 0.82; *p < 0.05 shown in blue; Pearson correlation). The aligned components (#1, #2, and #3) represent the generalized neural activity across sessions, while the non-correlated components (#4) represent session-specific neural activity. **b**, For cross-problem decoding, neurons recorded from different pairs of odor problems were

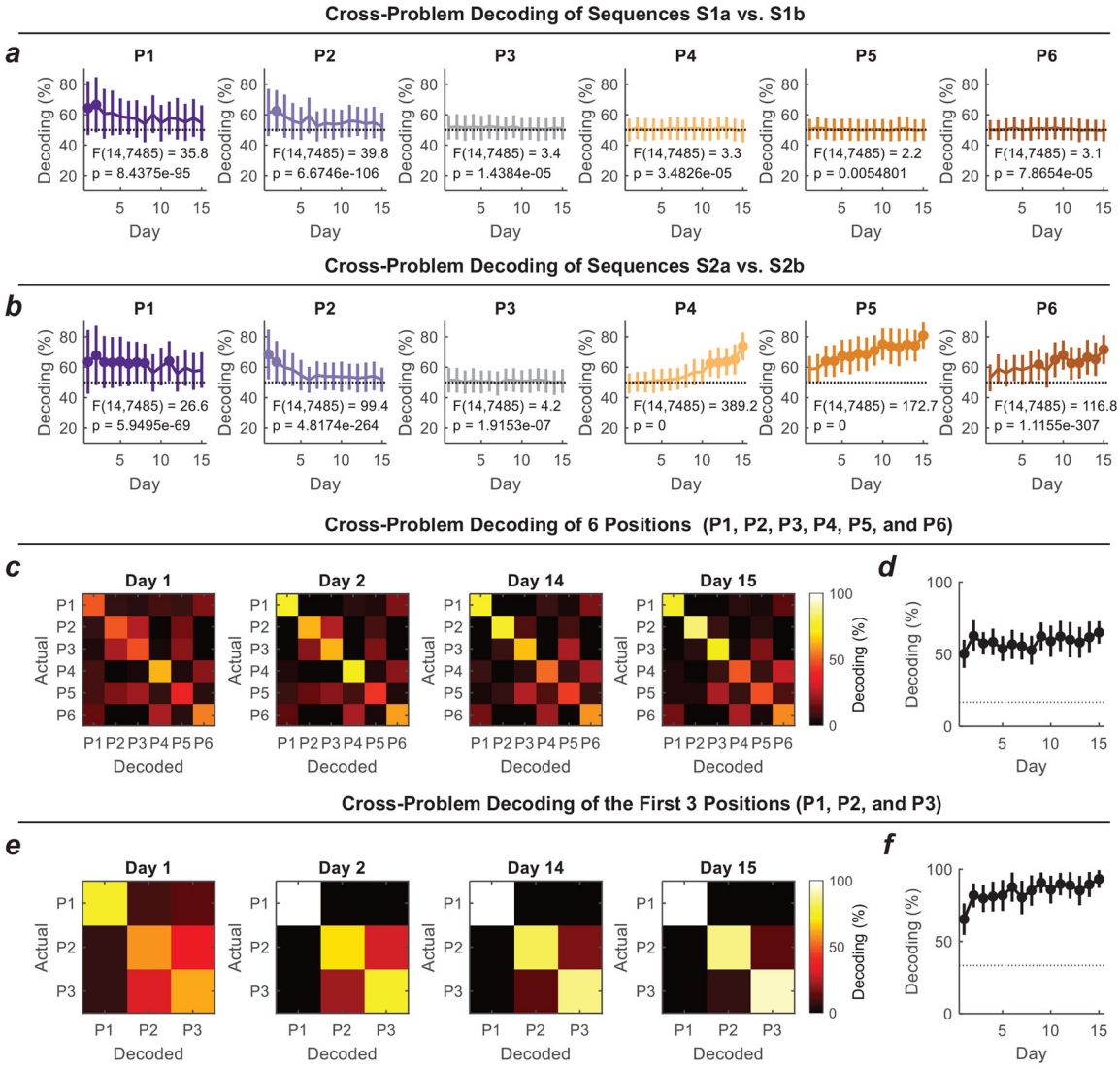
separately subjected to dimensionality reduction. The resulted two matrices (480 trials × 30 principal components; PCs) were aligned through CCA to obtain two correlated matrices (U1 and U2 for training; U3 and U4 for testing; 480 trials × 30 CCs for each matrix), which were concatenated (480 trials × 60 CCs) for further use as either a training (U1U2) or test (U3U4) set. Since there were 5 odor problems in total, for each repeat, the left one was combined with one of the two problems for the training set. For cross-subject decoding, the 9 rats were randomly separated into 4 groups: #1, #2, #3, and #4. Groups #1 and #2 were used to obtain the training set (U1U2), while Groups #3 and #4 were used to obtain the test set (U3U4).



Extended Data Fig. 6 |. Paired canonical components between problems.

a-b, This figure is an extension of Fig. 3a. Each panel plots one pair of canonical components (CCs), one from the training set and another from the test set, starting with CC4, the first CC not plotted in the main text figure. To obtain the training set, CCA was performed on a pair of problems (e.g. Problem #1 and Problem #2) to identify commonalities in the aligned neural subspaces. Similarly, to obtain the test set, the CCA was performed on a different pair of problems (e.g. Problem #3 and Problem #4). The scores of paired CCs (one from the training set and the other from the test set) were plotted against the 24 trial types for both Day 1 (**a**) and Day 15 (**b**). The 24 trial types are ordered as

P1(S1a,S1b,S2a,S2b), P2(S1a,S1b,S2a,S2b), P3(S1a,S1b,S2a,S2b), P4(S1a,S1b,S2a,S2b), P5(S1a,S1b,S2a,S2b), P6(S1a,S1b,S2a,S2b). Blue circles indicate positive trial types, and red circles indicate negative trial types. Four trial types are highlighted with filled circles (S2a4, S2b4, S2a5, and S2b5, in this order). The r in each panel is the correlation coefficient between paired CCs from the training and test sets (n = 480 trials for each CC; Pearson correlation). Data are presented as mean ± s.d.



Extended Data Fig. 7 |. Cross-problem decoding of sequences and positions.

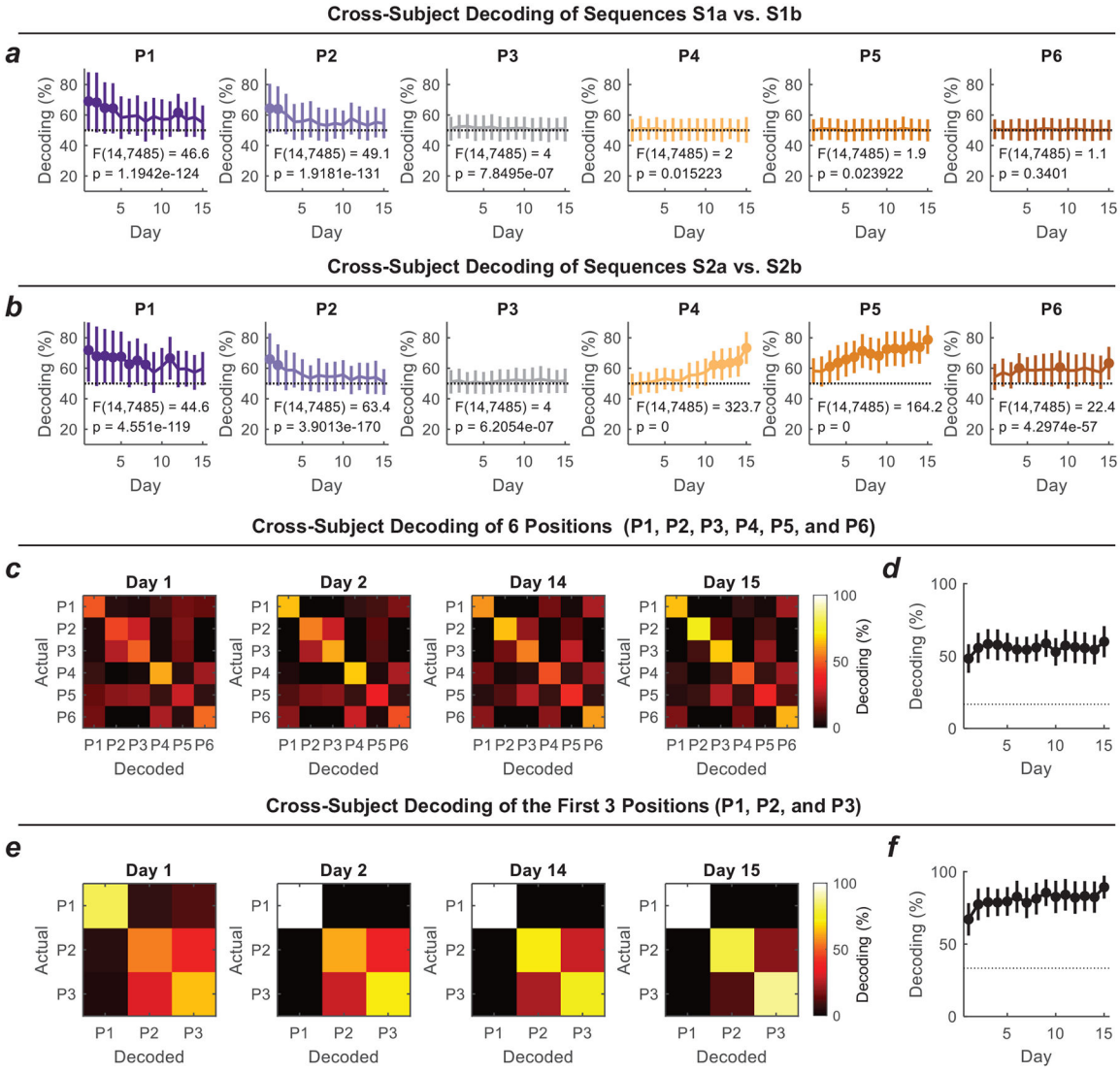
a, **b**, Cross-problem decoding of sequences S1a vs. S1b (**a**) and S2a vs. S2b (**b**) at six positions (P1 – P6). Training and test sets were generated with the same approach described in Extended Data Fig. 5b with all 480 trials. For decoding analysis in each panel, only particular trial types (e.g., S1a1 vs. S1b1) were selected. A one-way ANOVA (factor: Day) was performed and shown in each panel. **c**, Confusion matrices of 6 positions (P1 – P6) as a result of the cross-problem decoding of these positions. Trial types within the same position (e.g., S1a, S1b, S2a, and S2b) were lumped together. **d**, Cross-problem decoding of positions

during learning. A one-way ANOVA was used to test the effect of learning: $F(14,7485) = 85.4$, $p = 3\times10^{-228}$. **e**, Confusion matrices of the first three positions (P1 – P3) as a result of the cross-problem decoding of these three positions. Note that at each one of these three positions, the current and surrounding reward availabilities across sequences are similar, while at following positions (P4 and P5), the reward availabilities in S2b are not consistent with those in other sequences (S1a, S1b, S2a). **f**, Cross-problem decoding of the first three positions during learning. A one-way ANOVA was used to test the effect of learning: $F(14,7485) = 247.8$, $p = 0$. **a**, **b**, **d**, **f**, The round markers indicate that the mean decoding accuracy exceeded the 95% confidence interval (CI) of decoding accuracy from the same decoding process but with shuffled trial-type labels. Dotted line: chance level. n = 500 repeats. Data are presented as mean ± s.d.

Extended Data Fig. 8 |. Paired canonical components between subjects.

a-b, This figure is an extension of Fig. 4a. Each panel plots one pair of CCs, one from the training set and another from the test set, starting with CC #4, the first CC not plotted in the main text figure. For the training set, the CCA was performed on two groups of rats (e.g. rat group #1 and rat group #2) to identify commonalities in the aligned neural subspaces. For the test set, the CCA was performed on a different two groups of rats (e.g. rat group #3 and rat group #4). The scores of paired CCs (one from the training set and another from the test

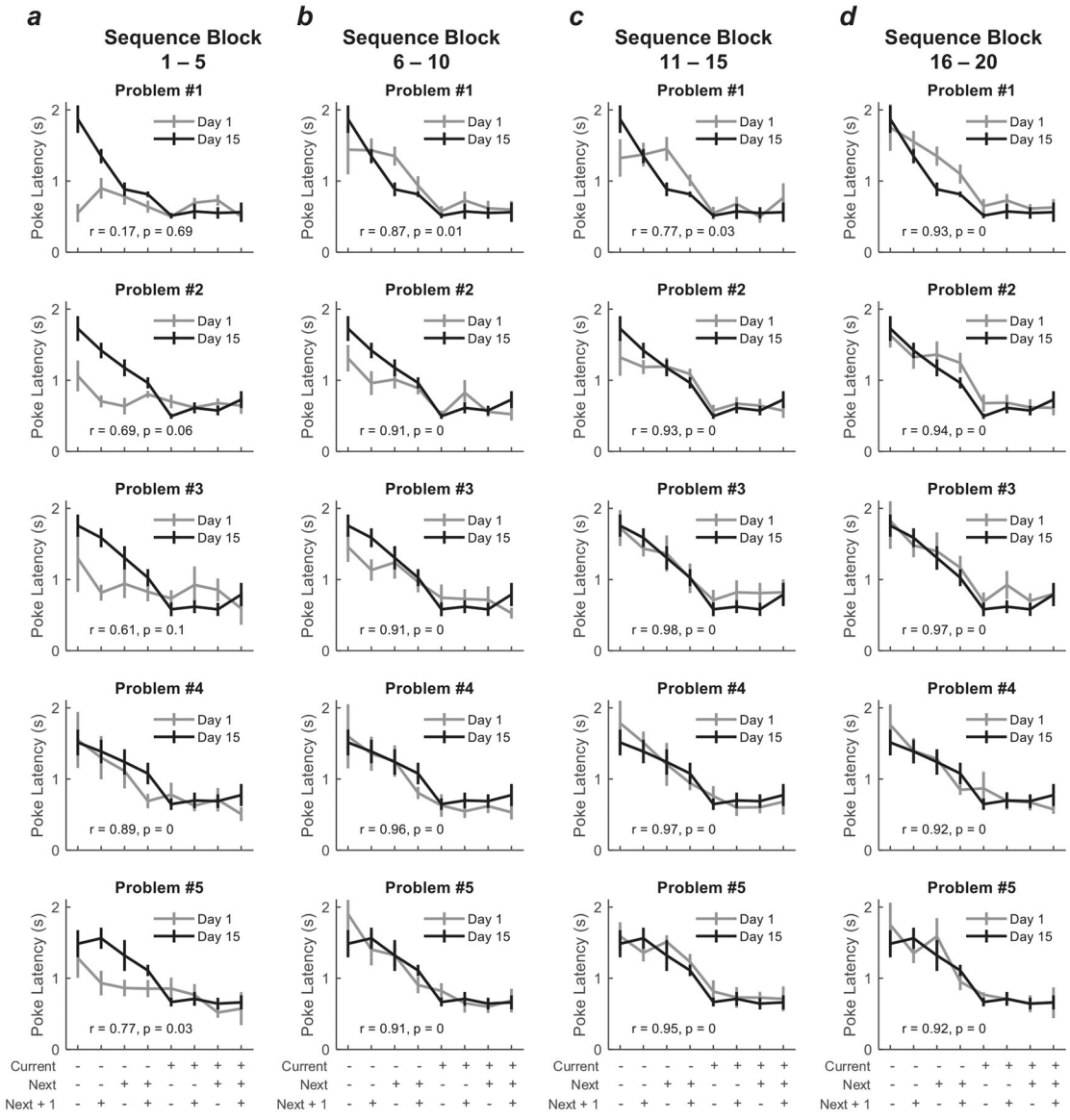
set) were plotted against the 24 trial types for both Day 1 (**a**) and Day 15 (**b**). The 24 trial types are ordered as P1(S1a,S1b,S2a,S2b), P2(S1a,S1b,S2a,S2b), P3(S1a,S1b,S2a,S2b), P4(S1a,S1b,S2a,S2b), P5(S1a,S1b,S2a,S2b), P6(S1a,S1b,S2a,S2b). Blue circles indicate positive trial types, and red circles indicate negative trial types. Four trial types were highlighted with filled circles (S2a4, S2b4, S2a5, and S2b5, in this order). The r in each panel is the correlation coefficient between paired CCs from the training and sets (n = 480 trials for each CC; Pearson correlation). Data are presented as mean ± s.d.



Extended Data Fig. 9 |. Cross-Subject Decoding of Sequences and Positions.

a, **b**, Cross-subject decoding of sequences S1a vs. S1b (**a**) and S2a vs. S2b (**b**) at six positions (P1 – P6). Training and test sets were generated with the same approach described in Extended Data Fig. 5b with all 480 trials. For decoding analysis in each panel, only particular trial types (e.g., S1a1 vs. S1b1) were selected. A one-way ANOVA (factor: Day) was performed and shown in each panel. **c**, Confusion matrices of 6 positions (P1 – P6) as a result of the cross-subject decoding of these positions. Trial types within the same position

(e.g., S1a, S1b, S2a, and S2b) were lumped together. **d**, Cross-subject decoding of positions during learning. A one-way ANOVA was used to test the effect of learning: $F(14,7485) = 38.1$, $p = 3.7\times10^{-101}$. **e**, Confusion matrices of the first three positions (P1 – P3) as a result of the cross-subject decoding of these three positions. Note that at each one of these three positions, the current and surrounding reward availabilities across sequences are similar, while at following positions (P4 and P5), the reward availabilities in S2b are not consistent with those in other sequences (S1a, S1b, S2a). **f**, Cross-subject decoding of the first three positions during learning. A one-way ANOVA was used to test the effect of learning: $F(14,7485) = 110.16$, $p = 3.1\times10^{-291}$. **a**, **b**, **d**, **f**, The round markers indicate that the mean decoding accuracy exceeded the 95% confidence interval (CI) of decoding accuracy from the same decoding process but with shuffled trial-type labels. Dotted line: chance level. n = 500 repeats. Data are presented as mean ± s.d.



Extended Data Fig. 10 |. Comparisons of poke latencies sorted by discounted future reward between Day 1 and Day 15.

a-d, Poke latency measuring the time from the onset of the houselights ("light") to nosepoke ("poke") at the odor port, sorted according to discounted future reward across **a**: sequence blocks 1 – 5; **b**, sequence blocks 6 – 10; **c**, sequence blocks 11 – 15; **d**, sequence blocks 16 – 20 on Day 1 (grey lines) compared to the poke latencies averaged across all 20 sequence blocks on Day 15 (black lines). Each sequence block has 24 trials and comes from 24 trial types. Analyses were performed on each odor problem separately (Problem #1 – #5; from upper to lower panels). In each panel, r is the correlation coefficient between the grey and black lines and p is the p-value for the correction. Number of sessions (i.e., rats) used in the analyses: Problem #1 (n = 9 on Day 1; n = 6 on Day 15); Problem #2 (n = 9 on Day 1; n = 8 on Day 15); Problem #3 (n = 7 on Day 1; n = 8 on Day 15); Problem #4 (n = 6 on Day 1; n = 7 on Day 15); Problem #5 (n = 6 on Day 1; n = 7 on Day 15). Data are presented as mean ± s.e.m.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank the NIDA IRP histology core for technical assistance with histology. This work used the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). This work was supported by a grant from the NIDA (K99DA049888 to J.Z.) and the Intramural Research Program at NIDA (ZIA-DA000587 to G.S.). The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

REFERENCES

1. Bartlett FC Remembering: a study in experimental and social psychology. (Cambridge Univ. Press, 1932).
2. Piaget J Langage et pensée chez l'enfant. (Delachaux et Niestlé, 1923).
3. van Kesteren MT, Ruiter DJ, Fernandez G & Henson RN How schema and novelty augment memory formation. Trends Neurosci 35, 211–219 (2012). [PubMed: 22398180]
4. Gilboa A & Marlatte H Neurobiology of schemas and schema-mediated memory. Trends Cogn Sci 21, 618–631 (2017). [PubMed: 28551107]
5. Tse D et al. Schemas and memory consolidation. Science 316, 76–82 (2007). [PubMed: 17412951]
6. Zhou J et al. Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of value and task structure in an odor sequence task. Curr Biol 29, 897–907 (2019). [PubMed: 30827919]
7. Zhou J et al. Complementary task structure representations in hippocampus and orbitofrontal cortex during an odor sequence task. Curr Biol 29, 3402–3409 (2019). [PubMed: 31588004]
8. Gallego JA, Perich MG, Chowdhury RH, Solla SA & Miller LE Long-term stability of cortical population dynamics underlying consistent behavior. Nat Neurosci 23, 260–270 (2020). [PubMed: 31907438]
9. Stringer C et al. Spontaneous behaviors drive multidimensional, brainwide activity. Science 364, 1–11 (2019).
10. Baram AB, Muller TH, Nili H, Garvert M, & Behrens TE Entorhinal and ventromedial prefrontal cortices abstract and generalise the structure of reinforcement learning problems. bioRxiv, 827253 (2020).
11. McKenzie S et al. Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. Neuron 83, 202–215 (2014). [PubMed: 24910078]

12. McKenzie S, Robinson NT, Herrera L, Churchill JC & Eichenbaum H Learning causes reorganization of neuronal firing patterns to represent related experiences within a hippocampal schema. J Neurosci 33, 10243–10256 (2013). [PubMed: 23785140]
13. Morrissey MD, Insel N & Takehara-Nishiuchi K Generalizable knowledge outweighs incidental details in prefrontal ensemble code over time. Elife 6 (2017).
14. Rubin A et al. Revealing neural correlates of behavior without behavioral measurements. Nat Commun 10, 4745 (2019). [PubMed: 31628322]
15. Mack ML, Preston AR & Love BC Ventromedial prefrontal cortex compression during concept learning. Nat Commun 11, 46 (2020). [PubMed: 31911628]
16. Farovik A et al. Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. J Neurosci 35, 8333–8344 (2015). [PubMed: 26019346]
17. Jones B & Mishkin M Limbic lesions and the problem of stimulus-reinforcement associations. Exp Neurol 36, 362–377 (1972). [PubMed: 4626489]
18. Jones JL et al. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science 338, 953–956 (2012). [PubMed: 23162000]
19. Wimmer GE & Shohamy D Preference by association: how memory mechanisms in the hippocampus bias decisions. Science 338, 270–273 (2012). [PubMed: 23066083]
20. Wilson RC, Takahashi YK, Schoenbaum G & Niv Y Orbitofrontal cortex as a cognitive map of task space. Neuron 81, 267–279 (2014). [PubMed: 24462094]
21. Constantinescu AO, O'Reilly JX & Behrens TEJ Organizing conceptual knowledge in humans with a gridlike code. Science 352, 1464–1468 (2016). [PubMed: 27313047]
22. Schuck NW, Cai MB, Wilson RC & Niv Y Human orbitofrontal cortex represents a cognitive map of state space. Neuron 91, 1402–1412 (2016). [PubMed: 27657452]
23. Garvert MM, Dolan RJ & Behrens TE A map of abstract relational knowledge in the human hippocampal-entorhinal cortex. Elife 6 (2017).
24. Behrens TEJ et al. What is a cognitive map? Oganizing kowledge for flexible behavior. Neuron 100, 490–509 (2018). [PubMed: 30359611]
25. Gardner MPH & Schoenbaum G The orbitofrontal cartographer. PsyArxiv (2020).
26. Gallagher M, McMahan RW & Schoenbaum G Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci 19, 6610–6614 (1999). [PubMed: 10414988]
27. Takahashi YK et al. Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. Neuron 80, 507–518 (2013). [PubMed: 24139047]
28. Stalnaker TA, Cooch NK & Schoenbaum G What the orbitofrontal cortex does not do. Nat Neurosci 18, 620–627 (2015). [PubMed: 25919962]
29. Schoenbaum G, Nugent SL, Saddoris MP & Setlow B Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport 13, 885–890 (2002). [PubMed: 11997707]
30. Gardner MPH, Conroy JS, Shaham MH, Styer CV & Schoenbaum G Lateral orbitofrontal inactivation dissociates devaluation-sensitive behavior and economic choice. Neuron 96, 1192–1203 (2017). [PubMed: 29154127]
31. Hirokawa J, Vaughan A, Masset P, Ott T & Kepecs A Frontal cortex neuron types categorically encode single decision variables. Nature 576, 446–451 (2019). [PubMed: 31801999]
32. Nogueira R et al. Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. Nat Commun 8, 14823 (2017). [PubMed: 28337990]
33. Young JJ & Shapiro ML Dynamic coding of goal-directed paths by orbital prefrontal cortex. J Neurosci 31, 5989–6000 (2011). [PubMed: 21508224]
34. Group ICA of fMRI toolbox (GIFT), <http://mialab.mrn.org/software/gift/> (2011).
35. Calhoun VD, Adali T, Pearlson GD & Pekar JJ A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp 14, 140–151 (2001). [PubMed: 11559959]
36. Hyvarinen A & Oja E Independent component analysis: algorithms and applications. Neural Netw 13, 411–430 (2000). [PubMed: 10946390]

37. McKeown MJ, Hansen LK & Sejnowsk TJ Independent component analysis of functional MRI: what is signal and what is noise? Curr Opin Neurobiol 13, 620–629 (2003). [PubMed: 14630228]
38. Wang J & Chang C-I Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. IEEE TGRS 44, 1586–1600 (2006).
39. Bell AJ & Sejnowski TJ An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7, 1129–1159 (1995). [PubMed: 7584893]
40. Long Q et al. Consistent run selection for independent component analysis: application to fMRI analysis. in IEEE ICASSP. 2581–2585 (2018).
41. Akhonda MABS, Levin-Schwartz Y, Bhinge S, Calhoun VD & Adali T Consecutive independence and correlation transform for multimodal fusion: application to EEG and fMRI data. in IEEE ICASSP. 2311–2315 (2018).
42. Jia C et al. C-ICT for discovery of multiple associations in multimodal imaging data: application to fusion of fMRI and DTI data. in CISS. 1–5 (2019).
43. Chang C-C & Lin C-J LIBSVM: A library for support vector machines. ACM TIST 2, 1–27 (2011).
44. Zhang Y et al. Object decoding with attention in inferior temporal cortex. Proc Natl Acad Sci U S A 108, 8850–8855 (2011). [PubMed: 21555594]

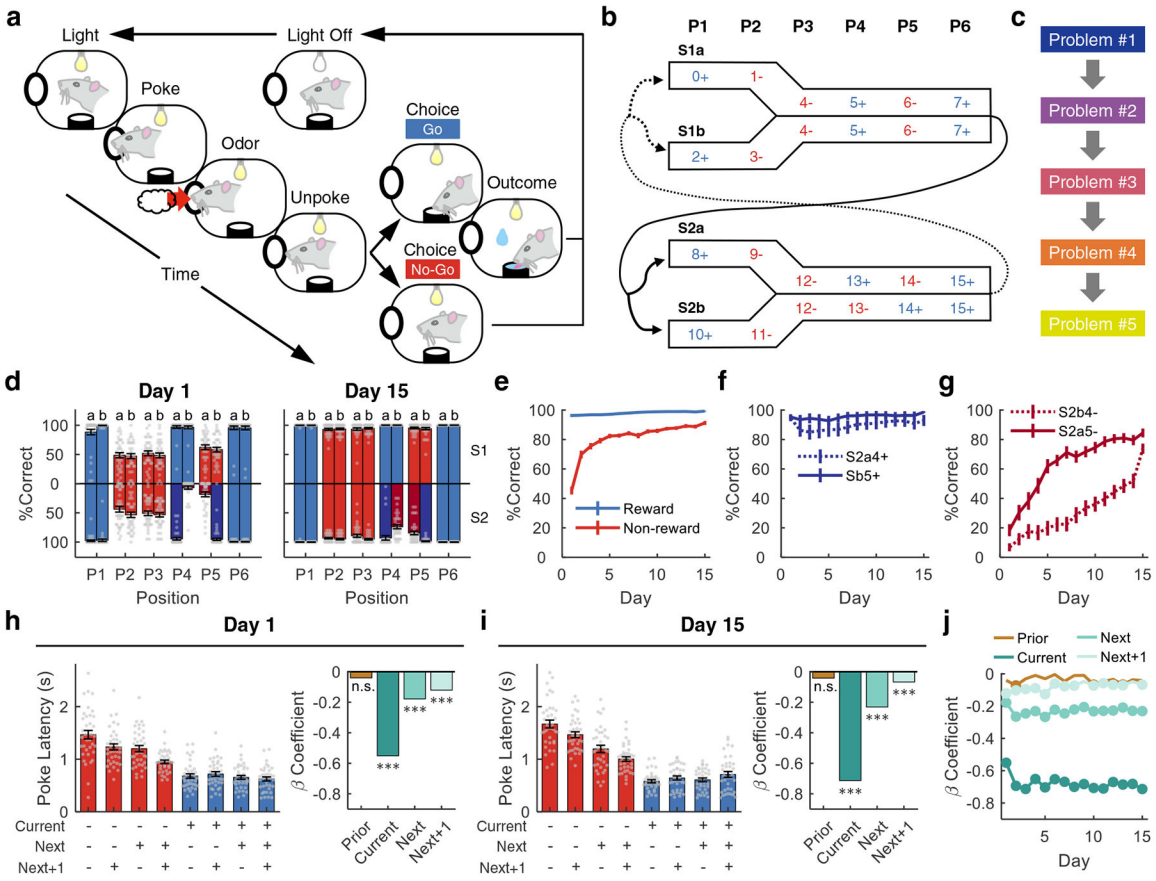


Fig. 1 |. Task design and behavior.

a, Single-trial illustration of the odor sequence task. **b**, 16 odors were organized into two pairs of sequences (S1a/b and S2a/b), each consisting of 6 trials or positions (P1 – P6). 0 – 15: odor identities. Blue/+: rewarded. Red/−: non-rewarded. Arrows: transitions between sequences. **c**, Each problem uses the same odor sequence structure but with 16 new odors. Nine rats were trained successively on five problems after shaping. Behavioral analyses combined the five problems and nine rats with data aligned to training days. **d**, Percent correct (%correct) in S1 (above) and S2 (below). Blue: rewarded; Red: non-rewarded. Darker colors: trial types that require rats to use past sequences. **e**, **f**, %Correct during learning. **h**, **i**, Poke latency (houselights onset to odor port entry) sorted by future rewards on Day 1 (**h**; left) and Day 15 (**i**; left). A linear regression for the poke latency was fit with four regressors (reward on the “prior”, “current”, “next”, and “next+1” trials) for Day 1 (**h**; right) and Day 15 (**i**; right). n.s., not significant; ***$p < 0.001$. **j**, Regression coefficients during learning. Round markers indicate statistical significance (*$p < 0.05$; corrected with the Benjamini-Hochberg procedure). **d**, **h**, **i**, n = 37 and 36 sessions for Day 1 and 15, respectively. **d-i**, Mean ± s.e.m. See Supplementary Table 3 for statistics.

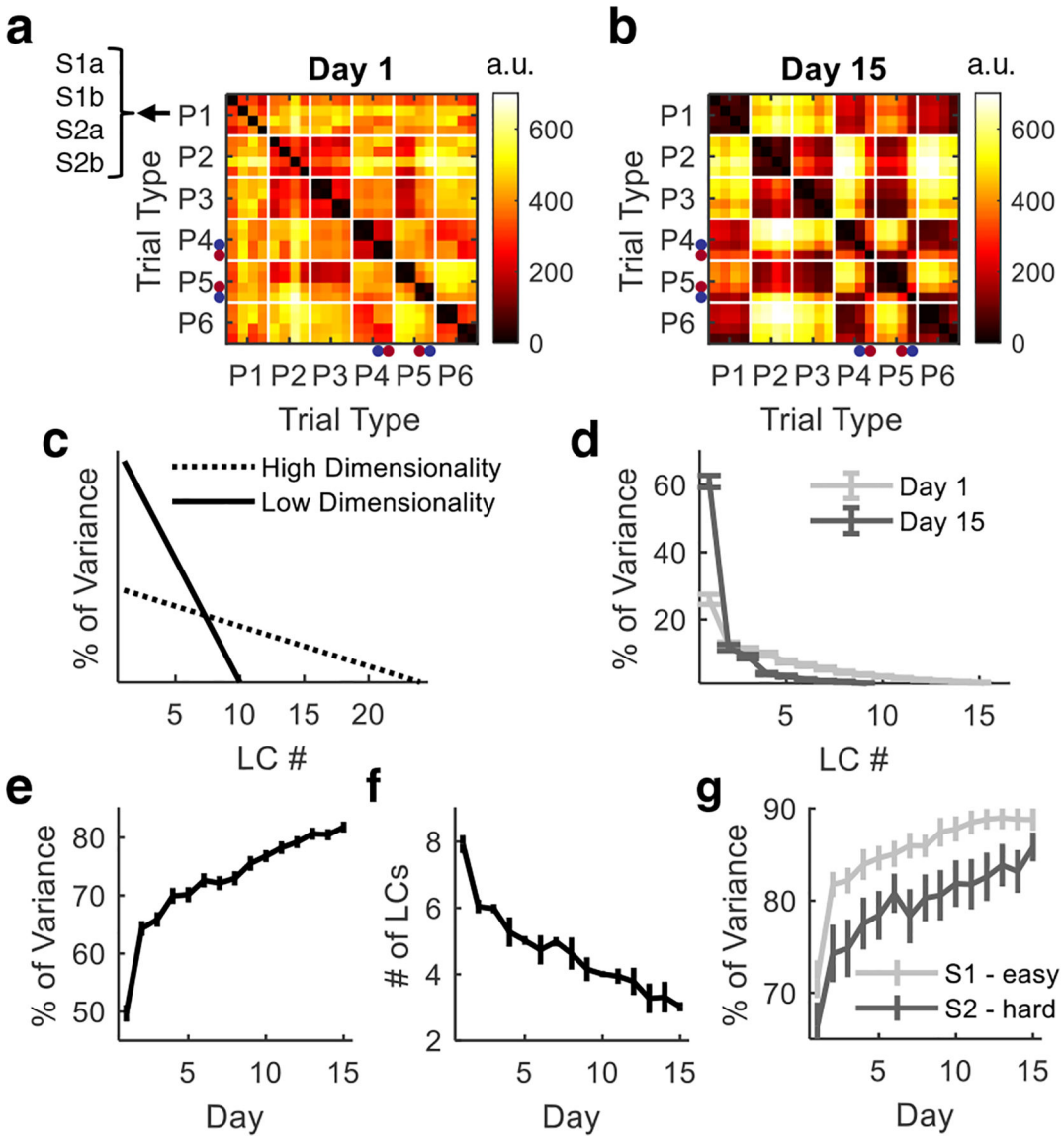


Fig. 2 |. Dimensionality of the task representation.

The neural data combined five odor problems and nine rats with alignment to the training days. For each repeat, the trial order within each one of the 24 trial types was shuffled to create a new pseudo-ensemble. An independent component analysis (ICA) was performed on the neural firing data to extract a low-dimensional representation of the neural data (480 trials × 100 neural dimensions), which then underwent a linear discriminant analysis (LDA) with the labels of 24 trial types. **a**, **b**, Dissimilarity matrices on Day 1 (**a**) and Day 15 (**b**) showing the Mahalanobis distance between each pair of trial types in the LDA activity space (averaged across 500 repeats). Within each position: S1a, S1b, S2a, and S2b. Blue dots: S2a4+ and S2b5+; red dots: S2a5− and S2b4−. **c**, Theoretical variance distributions. LC, linear discriminant component. **d**, Actual variance distributions on Day 1 and Day 15. **e**, The percent of variance explained by the first 3 LCs. **f**, The number of LCs needed to explain 80% variance of the data. **g**, The percent of variance explained by the first three LCs for

sequences S1 and S2 separately. **d-g**, n = 500 repeats. Mean ± s.d. See Supplementary Table 4 for statistics.

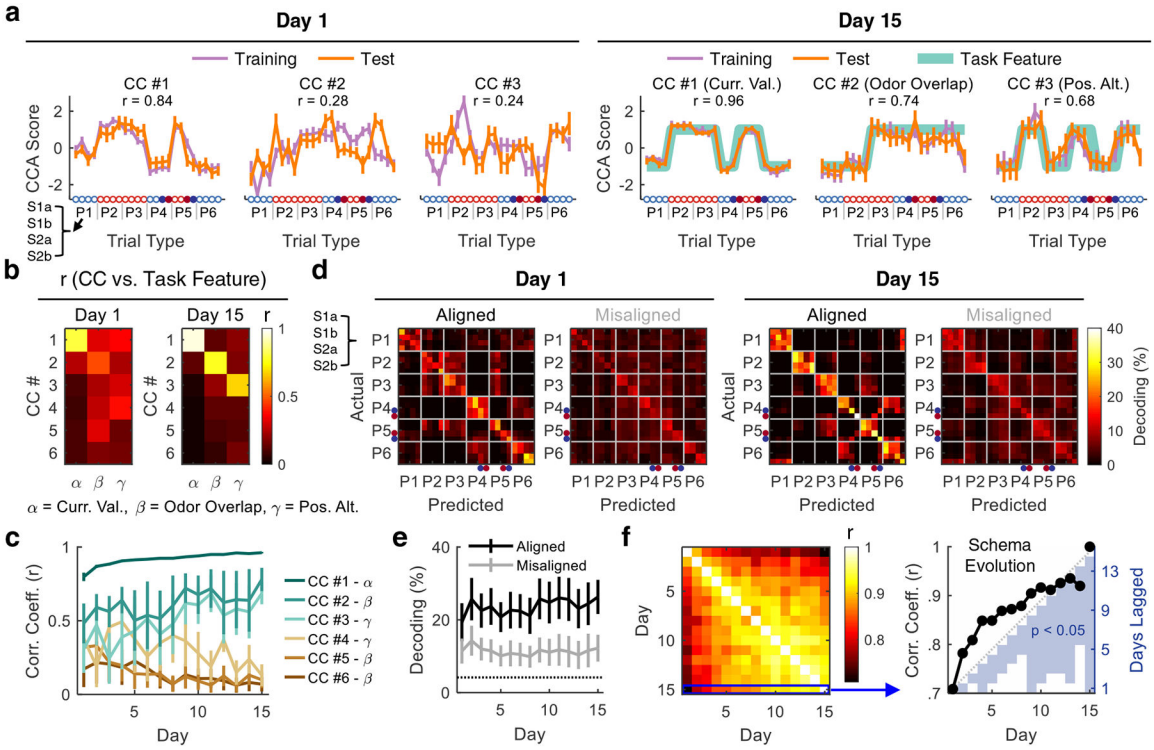


Fig. 3 |. Cross-problem decoding.

The training (480 trials × 60 CCs) and test (480 trials × 60 CCs) sets were separately obtained through dimensionality reduction and then manifold alignment on pseudo-ensembles from pairs of odor problems. CC, canonical component. **a**, The first three CCs in training and test from a single repeat. (r: correlation coefficient of paired CCs from training and test sets; n = 480 trials for each CC; Pearson correlation). Blue and red circles: rewarded and non-rewarded trials. Filled circles: S2a4+, S2b4−, S2a5−, and S2b5+. Curr. Val., Current Value; Pos. Alt., Positional Alternation. **b**, Correlation between the first 6 CCs and three task features. **c**, Correlation between the first 6 CCs and their matched task features. **d**, Decoding confusion matrices on Day 1 and Day 15. The trial-type orders for training and test sets were misaligned before manifold alignment in the control analysis (“Misaligned”). **e**, Comparing the mean decoding accuracy between the aligned and misaligned data. Dotted line: chance level. **f**, Left: correlation of confusion matrices between training days. Right: correlation between Day 15 and the other training days (y-axis, left). The blue patch indicates the correlation coefficient on one day (Day 2 – 15; x-axis) is significantly higher than that on another day that is n days lagged (n = 1 – 14; y-axis, right; *p < 0.05; one-tailed z-test after Fisher’s r-to-z transformation; corrected with the Benjamini-Hochberg procedure). **a-f**, n = 500 repeats. **a, c-e**, Mean ± s.d. See Supplementary Table 5 for statistics.

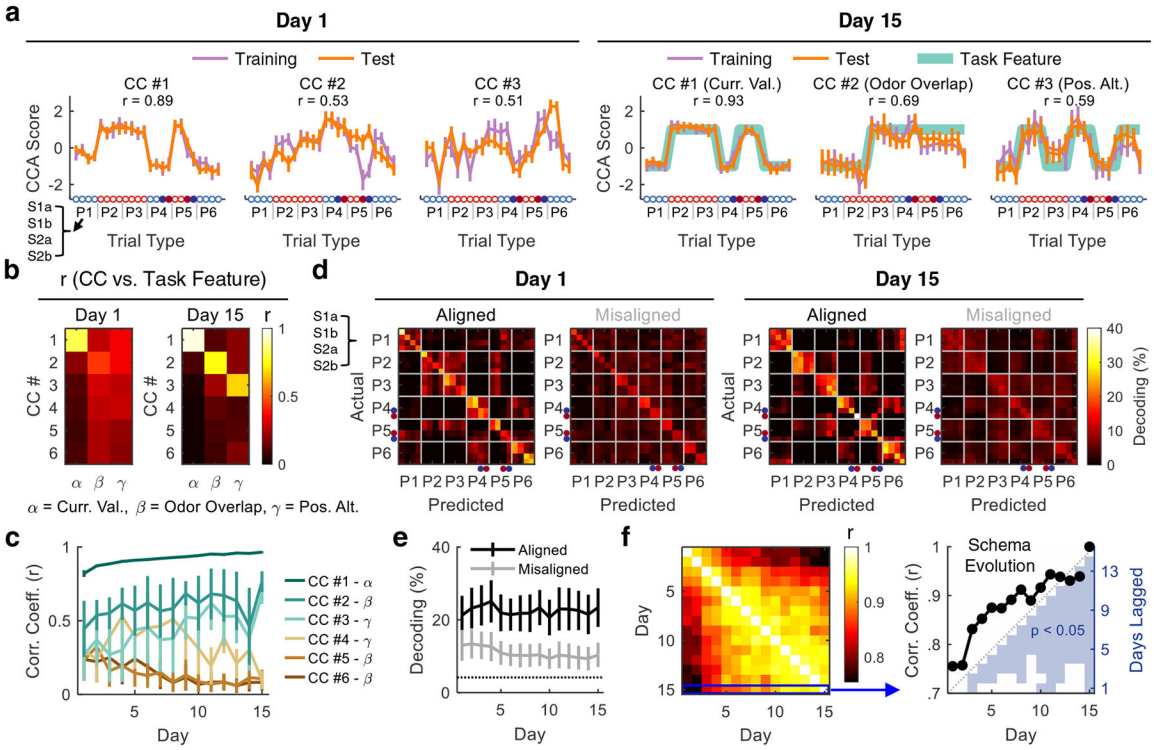


Fig. 4 |. Cross-subject decoding.

The description in Fig. 3 applies here except that training and test sets were separately obtained from pairs of rat groups rather than pairs of problems. For each repeat, the 9 rats were randomly separated into 4 groups: #1, #2, #3, and #4. Groups #1 and #2 were used for training, and #3 and #4 are used for testing. n = 500 repeats. Mean ± s.d. See Supplementary Table 6 for statistics.

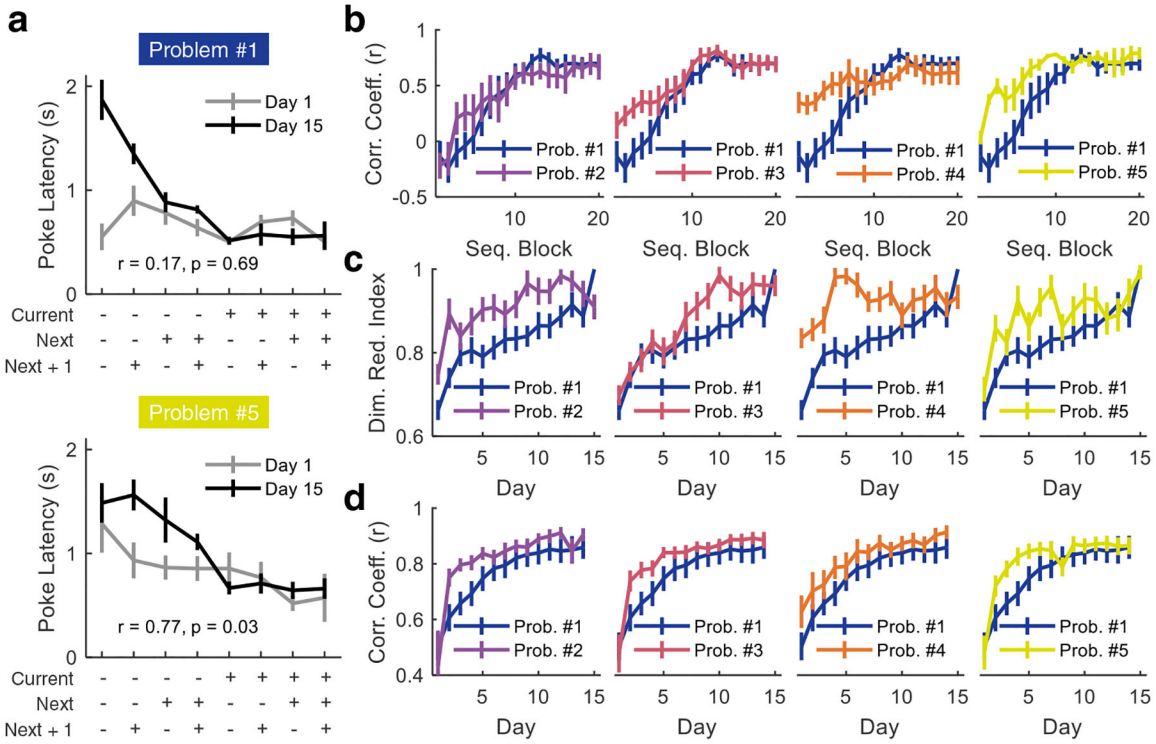


Fig. 5 |. Acceleration of behavioral and neural changes across problems.

a, Poke latencies for the first 5 sequence blocks on Day 1 (Problem #1: n = 9 rats; Problem #5: n = 6 rats) and all 20 sequence blocks on Day 15 (Problem #1: n = 6 rats; Problem #5: n = 6 rats), delineated by discounted future reward. Each block contains all 24 trial types. r and p (p-value): Pearson correlation between Day 1 and Day 15. Mean ± s.e.m. **b**, Correlation between poke latencies (sorted as in **a**) on Day 1 (5-block moving-window; step = 1 block) and Day 15 (fixed 20 blocks) for each of the 5 odor problems. Mean ± s.e.m. **c**, Neural dimensionality compression over time for each of the 5 odor problems (n = 500 repeats). Dim. Red. Index, Dimensionality Reduction Index, which was calculated as normalized percent of variance explained by the first three LCs. Mean ± s.d. **d**, Schema evolution measured how fast the confusion matrix obtained through cross-subject decoding analysis on any day became similar to that on Day 15 over training for each of the 5 odor problems (n = 500 repeats). Mean ± s.d. See Supplementary Table 7 for statistics.