A universal hippocampal memory code across animals and environments

Hannah S Wirtshafter[1*], Sara A Solla[1], John F Disterhoft[1]

*Lead contact: hsw@northwestern.edu (HSW)

1. Department of Neuroscience, Northwestern University Feinberg School of Medicine, Chicago, IL, USA.

How learning is affected by context is a fundamental question of neuroscience, as the ability to generalize learning to different contexts is necessary for navigating the world. An example of swift contextual generalization is observed in conditioning tasks, where performance is quickly generalized from one context to another. A key question in identifying the neural substrate underlying this ability is how the hippocampus (HPC) represents task-related stimuli across different environments, given that HPC cells exhibit place-specific activity that changes across contexts (remapping). In this study, we used calcium imaging to monitor hippocampal neuron activity as rats performed a conditioning task across multiple spatial contexts. We investigated whether hippocampal cells, which encode both spatial locations (place cells) and task-related information, could maintain their task representation even when their spatial encoding remapped in a new spatial context. To assess the consistency of task representations, we used advanced dimensionality reduction techniques combined with machine learning to develop manifold representations of population level HPC activity. The results showed that task-related neural representations remained stable even as place cell representations of spatial context changed, thus demonstrating similar embedding geometries of neural representations of the task across different spatial contexts. Notably, these patterns were not only consistent within the same animal across different contexts but also significantly similar across different animals, suggesting a standardized neural encoding or 'neural syntax' in the hippocampus. These findings bridge a critical gap between memory and navigation research, revealing how the hippocampus maintains cognitive consistency across different spatial environments. These findings also suggest that hippocampal function is governed by a neural framework shared between animals, an observation that may have broad implications for understanding memory, learning, and related cognitive processes. Looking ahead, this work opens new avenues for exploring the fundamental principles underlying hippocampal encoding strategies.

Introduction

How can learning can be generalized across contexts as well as remaining localized to one context? This question is fundamental to both neuroscience and philosophy[1-4]. Deficits in generalization or inappropriate generalization are hallmarks of many disorders, including autism[5,6], schizophrenia[7,8], and post-traumatic stress disorder[9,10]. In spite of their importance, many questions related to generalization remain to be answered.

The hippocampus (HPC) is important for learning, memory, and navigation, and damage to this region can disrupt contextual learning[11-17]. Many aspects of context, including an animal's spatial location and the presence of local and distal cues, can be represented by 'place cells' in the HPC[18-22]. The activity of many HPC cells therefore changes drastically in different environments (i.e. place cells remap), even when a task can be generalized across these different contexts[23-25]. A major open question is if and how representations of task-relevant stimuli, which are also found in the HPC[26-31], can be maintained against the background of remapped place cells. A further question is whether different animals solve this problem using the same, or similar, neural strategies.

Until recently, the neural mechanisms behind contextual learning have been challenging to investigate due to the need for tracking large numbers of cells across various environments and learning stages— tasks which were unachievable prior to the development of calcium imaging[32-37]. The use of calcium imaging in this study has allowed for a detailed exploration of hippocampal neuron dynamics, bridging previous gaps between studies of hippocampal

66 cognitive maps and the hippocampal bases of memory and learning. The traditional separation
67 between memory and navigation fields of HPC research is thus in the process of narrowing[38,39].

68 This research addresses two fundamental questions of learning and memory. The first
69 question pertains to the persistence of learning across varying contexts. The modulation of
70 hippocampal cells by spatial variables results in substantial changes to cell activity in different
71 environments[24,25,40,41]; how can conditioning-related neural representations remain stable amid
72 such remapping? The second question is the extent to which neural representations in the
73 hippocampus are invariant and consistent, not only within an individual across diverse contexts,
74 but across different individuals; is there a standardized neural encoding or 'neural syntax' in the
75 for learning and memory in the hippocampus? A commonality of encoding across animals would
76 imply that the functionality of the HPC is informed not solely by individual experiences but also
77 by a standardized framework of neural algorithms. Such a finding would provide key insights
78 into the underlying neural mechanisms that govern learning and memory, helping to identify
79 specific brain circuits and algorithms that drive behavior.

80 To answer these questions, we trained rats on an HPC-dependent conditioning task which is
81 rapidly generalized between spatial contexts[42,43]. We examined the same task in disparate
82 contexts and looked for changes in a representation of spatial location provided by a population
83 of place cells while the representation of the task features remains unchanged. We found that the
84 representations of the conditioning task were maintained as the animal generalized learning from
85 one environment to another, even as the representation of place changed. Surprisingly, we also
86 found that the neural representation of the task was consistent across animals.
87

This study demonstrates that despite the well-known phenomenon of place cell remapping[24,25,40,41], there exists a stable population-level neural representation of task features that persists across diverse environments. These representations also show consistency across different individuals, indicating a standardized neural encoding or 'neural syntax' for the conditioning task within the hippocampus. This novel finding suggests the existence of a universal coding mechanism for associative learning in the hippocampus, a principle that could reorient our approach to studying memory and could challenge our current understanding of cognitive processes.

Results

We trained a cohort of 5 freely moving rats in a conditioning task in one of two distinct environments, labeled A and B; the rats had been previously familiarized with the environment by the time the training began (Fig. 1a, Fig. S1). Environment A was an unscented rectangular enclosure with wire floor and walls, and white lighting; environment B was a scented ovular enclosure with white solid floor and walls, and red lighting (Fig. 1b). Both environments were located at the same spot in the room relative to external cues (see Methods); animals could see external cues out of the top of both environments, as well as out of the sides of environment A (animals also reared often, allowing them to see out of the sides of environment B). During training sessions, we recorded cellular activity in the hippocampal CA1 region via miniscopes, using Gcamp8m for calcium imaging (CaImg). To perform data analysis, we used both calcium events and calcium traces (see Methods), as indicated when applicable.

Animals easily transfer a conditioning task across environments

During the initial phase of our study, freely-moving rats (Fig. 1b) underwent training for trace eyeblink conditioning (tEBC) (Fig. S2), a hippocampus-dependent classical conditioning task that serves as a robust model for associative memory formation[44-47]. This paradigm involves presenting a 250ms conditioned stimulus (CS, in the form of a tone) followed by a 500ms trace interval, followed by the 100ms presentation of an unconditioned stimulus (US, an eyelid shock) (Fig. 1c). Both shock and blinking were recorded with wires inserted into the muscle of the eyelid (see Methods). As rats were trained, they exhibited a conditioned blink (CR) to the tone. Animals were considered to have learned the task after reaching criterion (70% CRs in 50 trials) on three consecutive training sessions (termed ‘criterion sessions’) or when the previous four

122 training sessions averaged over 70% (in this instance, only the final three of those sessions were
123 considered 'criterion sessions') (Fig. 1d). There was substantial variability in the number of
124 sessions it took to learn the task, for an average of 20±4.2 training sessions (note: number of
125 sessions always includes criterion sessions). The two rats that learned the fastest reached
126 criterion in 14 sessions, and the rat that learned the slowest reached criterion after 24 sessions
127 (Fig. S2). After reaching criterion, the rats were introduced to environment B, where their ability
128 to perform tEBC was assessed over a two-day period (one session per day). Comparative
129 analysis revealed no significant difference in performance (measured in % CRs) between the
130 criterion sessions in environment A and the testing phase in environment B (mean in
131 environment A criterion sessions was 74.75±6.49, mean in environment B test sessions was
132 77.70±11.68, two tailed t-test(24)=-0.83, $p>0.05$) (Fig. 1d), indicating the successful transfer of
133 tEBC learning to a new environment.
134
135 Calcium imaging (CaImg) enabled the longitudinal monitoring of the same hippocampal
136 cells over multiple sessions in both environments. For criterion and testing sessions, we observed
137 an average of 459.85±265.31 cells per session per animal, with no significant difference between
138 the number of cells recorded in environment A and environment B (two-tailed t-test(24)= -0.56,
139 $p>0.05$) On average, 132±95 cells were present in both the last criterion session in A and the first
140 testing session in B. This was not a significantly different numbers of cells that were present, on
141 average, in both the semi-final session in A, session A(n-1), and the final session in A, session
142 A(n) (155±115 cells).
143
144 ***Hippocampal place cell representations differ across environments***

To identify place cells, we compared actual mutual information (MI) to MI computed after bootstrapping circularly shifted position data 500 times. A cell was deemed a place cell if its MI was above the 95th percentile of null provided by the bootstrapped MI scores. Using this 95% cutoff, on average, 9.3%±4.2% of cells across criterion sessions in A and testing sessions in B qualified as place cells. There was not a significant difference between the percent of cells that qualified as place cells in environment A and environment B (average in environment A was 8.1%±3.3%, average in environment B was 11.1%±4.9%, two-tailed t-test $t(23) = -1.9$, $p>0.05$) (Fig. 2a, Fig. S3). The average mutual information scores across all criterion and testing sessions was 1.08±0.18, with no difference in average MI between environments (environment A mean MI was 1.08±0.17, environment B mean MI was 1.08±0.21, two-tailed t-test $t(4805)=-0.5$, $p>0.05$). (See methods for more detailed analysis).

Our analysis revealed that individual hippocampal (HPC) cells exhibited distinct spatial representations for environments A and B, altering their configurations of place cells and place fields relative to distal cues — a process known as 'place cell remapping' (Fig. 2b). We confirmed this remapping through several approaches. First, we quantified the shift in the location of highest calcium event rate (putative place field centers) by comparing their distances on the last two criterion sessions in environment A (sessions A(n-1) and A(n)) to the shift observed when transitioning from environment A on session n to environment B on session 1 (the centers of both environments were aligned, see methods). The data indicated a significantly greater change in these putative place field centers when the animals transitioned from A to B than when remaining within environment A (medians tested with Wilcoxon rank sum test $p=0.002$, means tested with double sided t-test $t(1430)= -2.5$, $p=0.01$, distributions tested with

168 two-sample Kolmogorov-Smirnov (KS) test, $p=3.6*10^{-4}$) (Fig. 2c). Second, to compensate for
169 any differences in environment size, we also compared the distances between field centers to the
170 distances expected if all centers were shuffled 100 times, and found that the median distance
171 between field centers when comparing session A(n) to session A(n-1) was less than all median
172 shuffled values (p=0), while the median distance between field centers when comparing session
173 A(n) to session B(1) was greater or equal than 56% of shuffled values (p=0.56) (Fig. 2d). Third,
174 accounting for the fact that cells may have multiple place fields, we computed the population
175 vector correlation (PVC[40,48,49], see methods) using calcium events for sessions A(n-1), A(n), and
176 B(1). When using cells that appeared in both sessions A(n-1) and A(n), we found a significant
177 positive correlation when computing the PVC for these two sessions (p = 0.0023, r = 0.11).
178 Conversely, when using cells that appeared in both sessions A(n) and B(1), we found no
179 correlation (p>0.05, r=-0.04). This result indicates significantly similar calcium event patterns
180 between sessions A(n-1) and session A(n), with no significant similarity in patterns between
181 session A(n) and B(1). (Fig. 2e).
182
183 We then used a machine learning algorithm to determine the variation in neural
184 embeddings between environments A and B. To do this, we applied the CEBRA algorithm[50] to
185 calcium trace imaging data labelled with spatial coordinates from environment A, session A(n)
186 (all were trained on 75% of data with 25% held out for verification). The choice of CEBRA was
187 motivated by its efficacy and interpretability in decoding neural activity patterns when compared
188 to alternative methods such as PCA[51] and Isomap[52] (see Methods for additional details). We then
189 tested this model's ability to decode the animals' position in environment A and environment B
190 when applied to neural data not used for training. The results showed that the model, when run

500 times, predicted the positions in session A(n-1) with significantly greater accuracy than what would be expected by chance; the null was constructed as determined when compared to a model trained on shuffled position data and also run 500x. (All double sided t-tests, for each rat: rat1: t(998):-34.3 $p=5.0*10^{-171}$, rat2: t(998):-72.5 p=0, rat3: t(998)=-1.96 p=0.05, rat4: t(998):-53.7 $p=2.6*10^{-299}$, rat5: t(998):-20.0 $p=4.1*10^{-75}$) (Fig. 3a-b). In contrast, when a model trained on data from A(n) was applied to environment B, the model's predictions were significantly below the accuracy of a model trained on shuffled position data, implying that the place cell coding across environments A and B are actually more different than would be expected by chance (all double sided t-tests, for each rat: rat1: t(998):84.1 p=0, rat2: t(998):18.8 $p=7.4*10^{-68}$, rat3: t(998)=74.7 p=0, rat4: t(998):13.1 $p=2.0*10^{-36}$, rat5: t(998):154.4 p=0) (Fig. 3c-d). Collectively, these findings suggest a significant remapping of place cells when transitioning between environments, and also that the neural embeddings for place coding in individual rats change when the animal switches contexts.

The hippocampus represents the conditioning task in both environments, and representations of the conditioning task are not spatial representations

We then investigated whether conditioning related data was represented equally in both environments A and B. It was obvious on visual inspection that individual cells varied their calcium event rate (Fig. 4a) and calcium trace (Fig. 4b) during the conditioning periods. To quantify this variation, we devised the metric 'CSUS mutual information' (CSUS-MI), analogous to spatial mutual information; this enabled us to assess the extent of task-related information captured by the calcium activity of each cell. We calculated the CSUS-MI for each cell and

214 benchmarked it against a control distribution generated by shuffling CS and US periods and
215 recalculating the MI 500 times. Using calcium event data, we found that 10.7±4.9% of cells held
216 significant CSUS information related to whether the animal was in a CS or US period (termed
217 CSUS-MI2 as the conditioning period was divided into 2 bins; see Methods) (Fig. S4a). An even
218 stronger relationship was noted if using calcium traces: 19.9±8.2 percent of cells contained
219 significant information related to whether the animal was in a CS or US period (Fig. S4a).
220 Importantly, neither of these MI metrics were significantly different between environments A
221 and B (double sided t-tests, using calcium events, $t(23)=0.48$, $p>0.05$, using calcium traces,
222 $t(23)=-0.52$, $p>0.05$).

223 We then extended this analysis to determine if calcium events or traces of individual cells
224 contained information about what temporal segment portion of the conditioning task the animal
225 was in. To do this, we divided the CSUS period into 5 equal sized bins, computed CSUS-MI
226 using these bins (termed CSUS-MI5, see methods), then compared these mutual information
227 values to the controls provided by shuffled data. Using calcium event data, we found that
228 15.5±7.8% of cells contained this information, compared to 10.0±7.8% of cells when we
229 calculated the MI using trace information (Fig. S4b). Again, neither of these mutual information
230 metrics were significantly different between environments A and B (double sided t-tests, using
231 calcium event data, $t(23)=-0.32$, using calcium trace data, $t(23)=-1.1$, $p>0.05$). There was not a
232 significant difference in CSUS-MI2 values when comparing values in session A(n) to session
233 A(n-1), versus comparing values in session A(n) to session B(1) (Wilcoxon rank sum test
234 $p>0.05$, double sided t-test, $t(1431)=0.86$, $p>0.05$). In contrast, there was a small but significant
235 difference in CSUS-MI5 when comparing session A(n) to session A(n-1) versus comparing A(n)
236 to session B(1) (Wilcoxon rank sum test $p=0.049$, double sided t-test, $t(1431)=-2.2$, $p=0.03$) (Fig.

237 S4c).Collectively, these results demonstrate that the conditioning task is represented in both
238 environments A and B, and that the percentage of cells representing the conditioning task was
239 not different between the two environments.

240 We then examined the overlap between cells that contained spatial information and those
241 with CSUS information, identifying a significant positive correlation between spatial MI and
242 CSUS-MI2 (linear regression, $r^2 = 0.04$, $p=2.1*10^{-104}$) and spatial MI and CSUS-MI5 (linear
243 regression, $r^2 = 0.09$, $p=1.7*10^{-237}$; note that the statistics have been computed for unbinned data
244 but the graph presents binned data for visualization purposes due to the large number of points)
245 (Fig. 4c). Further analysis revealed that cells with significant spatial modulation had a
246 significantly higher likelihood of being significantly modulated by CSUS compared to cells
247 without spatial modulation: 1.35 times higher chance of having significant spatial MI if the cell
248 has a significant CSUS-MI2 (Fisher's exact test, p=0.001) and 1.27 times higher chance if the
249 cell has a significant CSUS-MI5 (Fisher's exact test, p=0.002). We then inquired whether the
250 calcium events that occurred during conditioning periods were confined to the 'firing' fields of
251 place cells. We thus calculated the average location of calcium events during conditioning
252 periods versus the average location during periods of movement that were not conditioning
253 periods. We analyzed this data using the Mantel test, which statistically evaluates the correlation
254 between two distance matrices to determine if the spatial patterns they represent are significantly
255 related. Across all sessions, we found a Mantel statistic of 534.58; we compared this statistic to
256 the result of 10,000 shuffles to determine the statistic was not significant ($p>0.05$); i.e. the spatial
257 firing patterns during conditioning periods are not generally similar to those that occur during
258 non-conditioning periods (Fig. 4d). In other words, the spatial distribution of firing during

conditioning differs fundamentally from that during non-conditioning periods, a result which would not be expected if conditioning related responses were restricted to the cells' place fields

We next assessed the consistency of task representation across environments. When we compared cells that appeared in both sessions A(n) and B(1), there was no difference in cell responses to either CS or US (double sided t-tests for CS: $t(1174) = 0.68$ $p>0.05$ and ks test $p>0.05$, for US: $t(1174) = 1.20$ $p>0.05$, and ks-test $p>0.05$) (Fig. 4e).

Conditioning task representations are consistent across environments

We then trained a CEBRA model using calcium imaging data and time-stamped CS/US periods from environment A, using only cells that were recorded in both environment A and B. We then used this trained model to decode if the animal was in a CS or US period during an additional session in environment A, as well as in environment B. All models successfully decoded CS and US periods the additional session in environment A, as compared to shuffled data (all double sided t-tests, for each rat: rat1: $t(998):6.7$ $p=2.5*10^{-11}$, rat2: $t(998):16.1$ $p=7.4*10^{-52}$, rat3: $t(998)=83.5$ $p=0$, rat4: $t(998):80.1$ $p=0$, rat5: $t(998):61.0$ $p=0$) (Fig. 5a,c). All five models significantly outperformed chance level in environment B as determined by shuffled data (all double sided t-tests, for each rat: rat1: $t(998):10.4$ $p=2.6*10^{-24}$, rat2: $t(998):2.7$ $p=7.6*10^{-3}$, rat3: $t(998)=75.3$ $p=0$, rat4: $t(998):63.7$ $p=0$, rat5: $t(998):106.3$ $p=0$) (Fig. 5b-c).

We then trained an additional model on data from environment A during session A(n)to ascertain whether it could decode the temporal order within the conditioning period (CSUS5), both in an alternate session in environment A (session A(n-1)) and in environment B (session

B(1)). After being trained on session A(n), all models were able to decode environment A(n-1) well better than chance levels (all double sided t-tests for accuracy, for each rat: rat1: t(998):41.5 $p=3.3*10^{-220}$, rat2: t(998):28.7 $p=7.8*10^{-133}$, rat3: t(998)=122.6 p=0, rat4: t(998):118.5 p=0, rat5: t(998):62.6 p=0) (Fig. 5d,f,g). Remarkably, our results for decoding environment B(1) showed that temporal aspects of CS/US temporal order were decodable across environments, suggesting that a refined level of task encoding is stable across both environments (all double sided t-tests for accuracy, for each rat: rat1: t(998):55.1 $p=4.9*10^{-305}$, rat2: t(998):9.3 $p=1.1*10^{-19}$, rat3: t(998)=71.4 p=0, rat4: t(998):62.6 p=0, rat5: t(998):106.2 p=0; results were also significant compared to those for shuffled data for precision, recall, F1 score, and area under the receiver operating characteristic curve, data not shown, see Methods) (Fig. 5e-g).

Remarkably, for both CSUS2 and CSUS5, the model trained on session A(n) was no less accurate decoding session B(1) than it was decoding session A(n-1) (double sided t-tests, for CSUS2 t(8)=-0.13, p>0.05, for CSUS5, t(8)=0.32, p>0.05) (Fig. 5h).

We then used CEBRA to analyze the embedding geometries of cell representations during CS/US periods in both environments. First, we examined the embedding geometries for the conditioning task divided into CS and US periods (CSUS2), for 2, 3, 5, 7, and 10 latents. We compared sessions A(n-1), A(n), B(1), and B(2) to each other, as well as shuffled versions of each session.

For all 5 rats, the geometries displayed a high and significant degree of similarity as compared to the shuffled control. This significance was maintained when examining up to 10

304 latents, the largest number of latents we utilized (averaging results from all animals, all double
305 sided t-tests, for 2 latents all comparisons were $p<1*10^{-5}$, for 3 latents $p<1*10^{-4}$, for 5 latents
306 $p<1*10^{-5}$, for 7 latents $p<1*10^{-5}$, and for 10 latents $p<1*10^{-4}$) (Fig. 6a-c). This high degree of
307 similarity was maintained when the CS/US periods were divided into 5 segments (CSUS5, (all
308 double sided t-tests, for 2 latents all comparisons were $p<1*10^{-12}$, for 3 latents $p<1*10^{-7}$, for 5
309 latents $p<1*10^{-6}$, for 7 latents $p<1*10^{-5}$, and for 10 latents $p<1*10^{-5}$) (Fig. 6d-f).
310
311 These highly significant similarity signifies that the neural representations of the task
312 were consistent between environments A and B.
313

314 ***Conditioning task representations are consistent across animals***

315 Considering the similarity between representations of the task in environments A and B, we
316 wondered if there was a universal, inter-animal, representation of the conditioning task. To
317 answer this question, we investigated if there were coding similarities of the conditioning task
318 across subjects. For each animal, we developed a unique model based on calcium signal patterns
319 and the structure of the conditioning task. We then calculated a similarity score among all
320 animal-specific models. We observed a markedly significant consistency across these trained
321 models compared to those trained on shuffled data. Notably, this consistency was apparent in
322 models trained to differentiate between CS and US periods, as well as in more granular models
323 that recognized five discrete time segments during CS presentation, the trace interval, and US
324 delivery (akin to the models in Figures 6d-g). When the conditioning period is divided into 2
325 periods (CSUS2), the similarity across animal models is not significantly different than the

similarity between models in one animal for all tested number of latents (2 latents: ttest(188) = -0.45 $p>0.05$, 3 latents: ttest(188) = -1.57 $p>0.05$, 5 latents: ttest(188) = -0.40 $p>0.05$, 7 latents: ttest(188) = -0.22, $p>0.05$, 10 latents: ttest(188) = 0.40, $p>0.05$) (Fig. 7a-b). This relationship also holds when the conditioning period is divided into 5 periods (CSUS5): the similarity across animal models is not significantly different than the similarity between models in one animal (2 latents: ttest(188) = -0.79 $p>0.05$, 3 latents: ttest(188) = 0.25 $p>0.05$, 5 latents: ttest(188) = -0.42 $p>0.05$, 7 latents: ttest(188) = 0.70, $p>0.05$, 10 latents: ttest(188) = 0.30, $p>0.05$) (Fig. 7c-d).

Discussion

Our study provides significant insights into hippocampal function, as it demonstrates that the hippocampus not only responds to environmental change with changes in neural coding but also maintains consistent task-related information across varying contexts. These mechanisms underpin cognitive flexibility and the ability to apply learned behaviors in new situations. Below we will discuss this interplay between variability and consistency of representations through several theoretical lenses, including predictive coding and cognitive mapping theories, while exploring the stability of these processes within and across subjects.

Task abstraction across environments

A critical aspect of our study highlights that the hippocampus retains stable task representations, such as those required for eyeblink conditioning, despite variations in environmental contexts. This ability to generalize learned tasks across different settings supports models that posit a cognitive map that extends beyond simple spatial navigation, such as the Tolman-Eichenbaum Machine's (TEM)[53]. According to this model, "spatial" maps integrate task-related information and enable the hippocampus to utilize learned behaviors in novel environments that share cognitive demands but differ in sensory or environmental specifics. This flexible functionality exemplifies the hippocampus' role in abstracting and applying learned knowledge, a hallmark of high-dimensional cognitive mapping. This integrated framework facilitates the adaptation of learned behaviors across diverse contexts, an essential capability for navigating both physical and abstract environments.

Our findings further reveal that the hippocampus abstracts task-related information from the surrounding sensory environment and suggests that the hippocampus connects experiences

357 across different contexts by recognizing underlying similarities; these features help generalize
358 learning and adjust behavior. Our results thus support the TEM's perspective that the
359 hippocampus can encode higher-order, abstract information crucial for task execution. The
360 ability to detach task representation from immediate sensory inputs allows for a generalized
361 version of learned information[54-56], enhancing the hippocampus's utility in supporting the
362 organism's application of learned skills and behaviors in new albeit similar situations[55,56]. This
363 capacity for abstraction is indicative of a sophisticated neural coding mechanism and of an
364 adaptable and extensive cognitive mapping system, as it provides a buffer against potential
365 interference that could arise from the myriad of sensory stimuli an organism encounters. By
366 maintaining a conceptual, generalized version of learned information, the hippocampus supports
367 the organism's ability to apply learned skills and behaviors in new situations that share
368 underlying similarities with previous experiences but differ in sensory or contextual details.
369

370 **Pattern Separation vs. Completion**

371 Evidence from various areas of neuroscience has led to the development of a theory of HPC
372 function holding that the HPC treats states that involve equivalent actions or relationships such
373 as similar tasks as equivalent, resulting in learning that is easily transferred between
374 environments[53-55,57,58]. This theory contrasts with the theory that the HPC acts primarily to
375 perform competitive "pattern separation"[59-63]. The prevailing theory as well as the cellular and
376 systems level bases for contextual memory remain to be elucidated. In our study, the distinct
377 coding of different environments by hippocampal place cells provides evidence for pattern
378 separation: the hippocampus differentiates between distinct contexts. This separation reduces
379 interference between memories, allowing for more accurate recall based on specific

380 environmental cues. On the other hand, consistent decoding of the eyeblink task across different
381 environments suggests pattern completion. This process allows the hippocampus to reconstruct a
382 complete memory or learned response from partial or generalized cues, enabling the execution of
383 the learned task even when contextual details change.

384 Our results complement previous work that identifies the HPC as both a pattern
385 completer and a pattern separator. Studies have demonstrated that place cells can differentially
386 represent the same environment when task demands change[61,64,65], yet show similar firing
387 patterns when locations have similar task demands[66]. In more recent work[67], rats were exposed to
388 two distinct environments while performing variations of the same task: approaching object A in
389 the first environment and object B in the second; the study revealed anticorrelated hippocampal
390 firing patterns for events in the two contexts. This suggests that the hippocampus encodes
391 context-specific associations between items and locations, rather than just specific behavior. This
392 study underscored the role of the hippocampus in robust pattern separation when environments
393 differ but require similar behaviors, showing that even minor task variations can lead to
394 significant neuronal pattern separation[67]. In contrast, our study used a task that remained
395 identical across both environments and found consistent hippocampal population-level task
396 representations in both contexts. This consistency likely reflects that the task could be
397 generalized between environments, without necessitating the hippocampus to differentiate
398 between task demands. Therefore, the hippocampus seems to balance pattern separation and
399 completion based on how similar or distinct task demands are across different contexts.
400

401 **Non-spatial hippocampal representations**

There is ongoing debate about whether hippocampal pyramidal cells encode both spatial and non-spatial aspects of a context[39,68,69]. Our findings show that responses to the conditioning task were independent of place field location, with a fraction of the recorded individual cells found able to represent both spatial location and the conditioning task (Fig. 4). This contrasts with studies where conditioning responses were more closely tied to specific spatial locations[70,71]. The difference likely stems from task design: in previous studies, spatially contingent rewards or freezing behavior after a shock made location highly salient, by associating the place where the shock occurred with the aversive event. In our experiment, spatial position was irrelevant; this allowed pyramidal cells to encode task-relevant features independent of location. This finding aligns with previous work showing that hippocampal cells often respond to non-spatial aspects like sensory cues or task demands, particularly in non-spatial tasks[19,72-74].

The discovery that population level patterns in the hippocampus are organized into manifolds provides an elegant solution to the problem of single cells representing both spatial coordinates and task features. Previous hippocampal work has described distinct encoding for spatial location along the center stem of a T maze vs accumulation of evidence for a left-right turning decision at the end of this branch through two distinct, orthogonal directions in a two-dimensional neural manifold[75]. Other brain regions, such as the prefrontal cortex and cingulate cortex, appear to use a similar orthogonal coding strategy[76-78]; this strategy has also emerged in neural network simulations of a context-dependent classification task[79].

Intra-Subject and Inter-Subject Consistency

The stability of neural representations within subjects across different testing sessions indicates that once hippocampal circuits are trained, their functional architecture remains remarkably consistent, even in varying contexts. This intra-subject consistency supports theories suggesting that neural circuits are not just reactive but possess a robust, predefined role in processing and responding to specific stimuli[80,81]. Furthermore, the observation of similar neural encoding patterns across different animals performing the same task suggests a species specific, possibly evolutionary conserved, neural code[80,81]. These findings highlight a generalized neural processing strategy that may have been shaped by natural selection to optimize cognitive and behavioral responses across environmental challenges faced by a species. Such a generalized coding strategy may be indicative of evolutionary pressures that have favored neural mechanisms promoting cognitive flexibility and rapid adaptation to environmental challenges[82-84].

The results of our study highlight the preservation of hippocampal task encoding across different contexts and species; this presents a surprising parallel to recent findings in motor cortex[85,86] and insular cortex[87]. Motor functions, especially those fundamental to survival and interaction with the physical environment, are expected to exhibit conserved neural dynamics due to their innate and reflexive nature; motor tasks typically involve stereotyped and predictable patterns of behavior that are essential for immediate responses and interactions with the environment[88-91]. Similarly, the motivational states preserved across animals in the insular cortex are those central to basic biological needs, such as thirst and hunger. These states and tasks are often highly conserved across individuals because they rely on well-established neural circuits that perform specific, crucial functions necessary for survival[89,91-95].

In contrast, hippocampal tasks involve complex cognitive processes that include memory, learning, and spatial navigation; these require a higher degree of cognitive flexibility[83,96-98]. and

are generally influenced by individual experiences (and, at the species level, specific ecological and evolutionary pressures)[99-103]. Therefore, the conservation of hippocampal task encoding across different individuals, as observed in our study, challenges these traditional views and suggests a deeper, possibly adaptive significance to these cognitive functions.

The surprising conservation of the neural representation of tasks within the hippocampus suggests that certain aspects of cognitive mapping and memory processing might be as evolutionarily essential as motor functions. This conservation might reflect universal cognitive strategies that are critical for survival across a range of environmental contexts, providing individuals within a species with the ability to adapt behavior based on past experiences and anticipated future conditions. Such a mechanism would not only enhance an organism's ability to navigate complex environments but also facilitate learning and decision-making across generational timescales.

Conclusion

The consistent decoding of eyeblink conditioning tasks across different environmental contexts indicates that the hippocampus can maintain a stable representation of task-specific information irrespective of the external sensory environment. This suggests an advanced capability for abstract cognitive mapping, where the hippocampus constructs and utilizes cognitive maps not only for physical locations but also for abstract tasks and concepts, allowing for effective application in varying contexts. These findings expand our understanding of how memories are formed, stored, and retrieved. They suggest that memories are not just static

recollections of past events but dynamic and adaptable representations that can be applied to new situations.

Citations

1 Bouton, M. E. A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health psychology* **19**, 57 (2000).

2 Herszage, J. & Censor, N. Modulation of learning and memory: a shared framework for interference and generalization. *Neuroscience* **392**, 270-280 (2018).

3 Linda, Q. Y., Wilson, R. C. & Nassar, M. R. Adaptive learning is structure learning in time. *Neuroscience & Biobehavioral Reviews* **128**, 270-281 (2021).

4 Censor, N. Generalization of perceptual and motor learning: a causal link with memory encoding and consolidation? *Neuroscience* **250**, 201-207 (2013).

5 Brown, S. & Bebko, J. Generalization, overselectivity, and discrimination in the autism phenotype: A review. *Research in Autism Spectrum Disorders* **6**, 733-740 (2012).

6 Church, B. A. *et al.* Learning, plasticity, and atypical generalization in children with autism. *Psychonomic Bulletin & Review* **22**, 1342-1348 (2015).

7 Weiler, J. A., Bellebaum, C., Brüne, M., Juckel, G. & Daum, I. Impairment of probabilistic reward-based learning in schizophrenia. *Neuropsychology* **23**, 571 (2009).

8 Shohamy, D. *et al.* Learning and generalization in schizophrenia: effects of disease and antipsychotic drug treatment. *Biological psychiatry* **67**, 926-932 (2010).

9 Anastasides, N. *et al.* Increased generalization of learned associations is related to re-experiencing symptoms in veterans with symptoms of post-traumatic stress. *Stress* **18**, 484-489 (2015).

10 Lis, S. *et al.* Generalization of fear in post-traumatic stress disorder. *Psychophysiology* **57**, e13422 (2020).

11 Calandreau, L., Desgranges, B., Jaffard, R. & Desmedt, A. Switching from contextual to tone fear conditioning and vice versa: the key role of the glutamatergic hippocampal-lateral septal neurotransmission. *Learn Mem* **17**, 440-443 (2010). https://doi.org:10.1101/lm.1859810

12 Desmedt, A., Garcia, R. & Jaffard, R. Vasopressin in the lateral septum promotes elemental conditioning to the detriment of contextual fear conditioning in mice. *Eur J Neurosci* **11**, 3913-3921 (1999).

13 Vouimba, R. M., Garcia, R. & Jaffard, R. Opposite effects of lateral septal LTP and lateral septal lesions on contextual fear conditioning in mice. *Behav Neurosci* **112**, 875-884 (1998).

14 Anagnostaras, S. G., Gale, G. D. & Fanselow, M. S. Hippocampus and contextual fear conditioning: recent controversies and advances. *Hippocampus* **11**, 8-17 (2001). https://doi.org:10.1002/1098-1063(2001)11:1<8::AID-HIPO1015>3.0.CO;2-7

15 Honey, R. C. & Good, M. Selective hippocampal lesions abolish the contextual specificity of latent inhibition and conditioning. *Behav Neurosci* **107**, 23-33 (1993). https://doi.org:10.1037//0735-7044.107.1.23

16 Plitt, M. H. & Giocomo, L. M. Experience-dependent contextual codes in the hippocampus. *Nat Neurosci* **24**, 705-714 (2021). https://doi.org:10.1038/s41593-021-00816-6

17 Devito, L. M., Kanter, B. R. & Eichenbaum, H. The hippocampus contributes to memory expression during transitive inference in mice. *Hippocampus* **20**, 208-217 (2010). https://doi.org:10.1002/hipo.20610

18 Schimanski, L. A., Lipa, P. & Barnes, C. A. Tracking the course of hippocampal representations during learning: when is the map required? *J Neurosci* **33**, 3094-3106 (2013). https://doi.org:10.1523/JNEUROSCI.1348-12.2013

19 Gauthier, J. L. & Tank, D. W. A Dedicated Population for Reward Coding in the Hippocampus. *Neuron* **99**, 179-193 e177 (2018). https://doi.org:10.1016/j.neuron.2018.06.008

20 Mamad, O. *et al.* Place field assembly distribution encodes preferred locations. *PLoS Biol* **15**, e2002365 (2017). https://doi.org:10.1371/journal.pbio.2002365

21 Hok, V. *et al.* Goal-related activity in hippocampal place cells. *J Neurosci* **27**, 472-482 (2007). https://doi.org:27/3/472 [pii] 10.1523/JNEUROSCI.2864-06.2007

22 Knierim, J. J. Dynamic interactions between local surface cues, distal landmarks, and intrinsic circuitry in hippocampal place cells. *J Neurosci* **22**, 6254-6264 (2002). https://doi.org:20026608

23 Allen, K., Rawlins, J. N., Bannerman, D. M. & Csicsvari, J. Hippocampal place cells can encode multiple trial-dependent features through rate remapping. *J Neurosci* **32**, 14752-14766 (2012). https://doi.org:10.1523/JNEUROSCI.6175-11.2012

24 Jeffery, K. J., Gilbert, A., Burton, S. & Strudwick, A. Preserved performance in a hippocampal-dependent spatial task despite complete place cell remapping. *Hippocampus* **13**, 175-189 (2003). https://doi.org:10.1002/hipo.10047

25 Colgin, L. L., Moser, E.I., May-Britt Moser. Understanding memory through hippocampal remapping. *Cell* (2008).

26 Christian, K. M. & Thompson, R. F. Neural substrates of eyeblink conditioning: acquisition and retention. *Learn Mem* **10**, 427-455 (2003). https://doi.org:10.1101/lm.59603

27 McEchron, M. D., Tseng, W. & Disterhoft, J. F. Single neurons in CA1 hippocampus encode trace interval duration during trace heart rate (fear) conditioning in rabbit. *J Neurosci* **23**, 1535-1547 (2003).

28 Wirtshafter, H. S. & Wilson, M. A. Locomotor and Hippocampal Processing Converge in the Lateral Septum. *Current biology : CB* **29**, 3177-3192 (2019). https://doi.org:10.1016/j.cub.2019.07.089

29 Berger, T. W., Rinaldi, P. C., Weisz, D. J. & Thompson, R. F. Single-unit analysis of different hippocampal cell types during classical conditioning of rabbit nictitating membrane response. *J Neurophysiol* **50**, 1197-1219 (1983). https://doi.org:10.1152/jn.1983.50.5.1197

30 Wirtshafter, H. S. & Wilson, M. A. Differences in reward biased spatial representations in the lateral septum and hippocampus. *eLife* **9** (2020). https://doi.org:10.7554/eLife.55252

31 Wirtshafter, H. S. & Disterhoft, J. F. Place cells are nonrandomly clustered by field location in CA1 hippocampus. *Hippocampus* **33**, 65-84 (2023). https://doi.org:10.1002/hipo.23489

32 Aharoni, D. & Hoogland, T. M. Circuit Investigations With Open-Source Miniaturized Microscopes: Past, Present and Future. *Frontiers in cellular neuroscience* **13**, 141 (2019). https://doi.org:10.3389/fncel.2019.00141

33 Aharoni, D., Khakh, B. S., Silva, A. J. & Golshani, P. All the light that we can see: a new era in miniaturized microscopy. *Nat Methods* **16**, 11-13 (2019). https://doi.org:10.1038/s41592-018-0266-x

34 Silva, A. J. Miniaturized two-photon microscope: seeing clearer and deeper into the brain. *Light: Science & Applications* **6**, e17104-e17104 (2017).

35 Wirtshafter, H. S. & Disterhoft, J. F. In Vivo Multi-Day Calcium Imaging of CA1 Hippocampus in Freely Moving Rats Reveals a High Preponderance of Place Cells with Consistent Place Fields. *J Neurosci* (2022). https://doi.org:doi: 10.1523/JNEUROSCI.1750-21.2022

36 Cunningham, J. P. & Byron, M. Y. Dimensionality reduction for large-scale neural recordings. *Nature neuroscience* **17**, 1500-1509 (2014).

37 Bassett, D. S. & Sporns, O. Network neuroscience. *Nature neuroscience* **20**, 353-364 (2017).

38 Ekstrom, A. D. & Ranganath, C. Space, time, and episodic memory: The hippocampus is all over the cognitive map. *Hippocampus* **28**, 680-687 (2018).

39 Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M. & Tanila, H. The hippocampus, memory, and place cells: is it spatial memory or a memory space? *Neuron* **23**, 209-226 (1999). https://doi.org:S0896-6273(00)80773-4 [pii]

40 Fenton, A. A. Remapping revisited: how the hippocampus represents different spaces. *Nat Rev Neurosci* **25**, 428-448 (2024). https://doi.org:10.1038/s41583-024-00817-x

41 Sanders, H., Wilson, M. A. & Gershman, S. J. Hippocampal remapping as hidden state inference. *eLife* **9** (2020). https://doi.org:10.7554/eLife.51140

42 Bouton, M. E., Nelson, J. B. & Rosas, J. M. Stimulus generalization, context change, and forgetting. *Psychological bulletin* **125**, 171-186 (1999). https://doi.org:10.1037/0033-2909.125.2.171

43 Lovibond, P. F., Preston, G. & Mackintosh, N. Context specificity of conditioning, extinction, and latent inhibition. *Journal of Experimental Psychology: Animal Behavior Processes* **10**, 360 (1984).

44 Moyer, J. R., Jr., Deyo, R. A. & Disterhoft, J. F. Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behav Neurosci* **104**, 243-252 (1990).

45 McEchron, M. D. & Disterhoft, J. F. Hippocampal encoding of non-spatial trace conditioning. *Hippocampus* **9**, 385-396 (1999). https://doi.org:10.1002/(SICI)1098-1063(1999)9:4<385::AID-HIPO5>3.0.CO;2-K

46 Woodruff-Pak, D. S. & Disterhoft, J. F. Where is the trace in trace conditioning? *Trends Neurosci* **31**, 105-112 (2008). https://doi.org:10.1016/j.tins.2007.11.006

47 Weiss, C. & Disterhoft, J. F. The impact of hippocampal lesions on trace-eyeblink conditioning and forebrain-cerebellar interactions. *Behav Neurosci* **129**, 512-522 (2015). https://doi.org:10.1037/bne0000061

48 Neymotin, S. A., Talbot, Z. N., Jung, J. Q., Fenton, A. A. & Lytton, W. W. Tracking recurrence of correlation structure in neuronal recordings. *J Neurosci Methods* **275**, 1-9 (2017). https://doi.org:10.1016/j.jneumeth.2016.10.009

49 Brun, V. H. *et al.* Progressive increase in grid scale from dorsal to ventral medial entorhinal cortex. *Hippocampus* **18**, 1200-1212 (2008). https://doi.org:10.1002/hipo.20504

50 Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360-368 (2023). https://doi.org:10.1038/s41586-023-06031-6

51 Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science* **2**, 559-572 (1901).

52 Tenenbaum, J. B., Silva, V. d. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *science* **290**, 2319-2323 (2000).

53 Whittington, J. C. R. *et al.* The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell* **183**, 1249-1263 e1223 (2020). https://doi.org:10.1016/j.cell.2020.10.024

54 Wirtshafter, H. S. & Wilson, M. A. Artificial intelligence insights into hippocampal processing. *Front Comput Neurosci* **16**, 1044659 (2022). https://doi.org:10.3389/fncom.2022.1044659

55 Lehnert, L., Littman, M. L. & Frank, M. J. Reward-predictive representations generalize across tasks in reinforcement learning. *PLoS Comput Biol* **16**, e1008317 (2020). https://doi.org:10.1371/journal.pcbi.1008317

56 Collins, A. G. E. & Frank, M. J. Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* **152**, 160-169 (2016).

57 Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map. *Nat Neurosci* **20**, 1643-1653 (2017). https://doi.org:10.1038/nn.4650

58 Eichenbaum, H. Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* **44**, 109-120 (2004).

59 Wills, T. J., Lever, C., Cacucci, F., Burgess, N. & O'Keefe, J. Attractor dynamics in the hippocampal representation of the local environment. *Science* **308**, 873-876 (2005). https://doi.org:10.1126/science.1108905

60 Colgin, L. L. *et al.* Attractor-map versus autoassociation based attractor dynamics in the hippocampal network. *J Neurophysiol* **104**, 35-50 (2010). https://doi.org:10.1152/jn.00202.2010

61 Smith, D. M. & Mizumori, S. J. Learning-related development of context-specific neuronal responses to places and events: the hippocampal role in context processing. *J Neurosci* **26**, 3154-3163 (2006). https://doi.org:10.1523/JNEUROSCI.3234-05.2006

62 Colgin, L. L. *et al.* Attractor-map versus autoassociation based attractor dynamics in the hippocampal network. *Journal of neurophysiology* **104**, 35-50 (2010).

63 Lee, I., Yoganarasimha, D., Rao, G. & Knierim, J. J. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature* **430**, 456-459 (2004).

64 Bahar, A. S., Shirvalkar, P. R. & Shapiro, M. L. Memory-guided learning: CA1 and CA3 neuronal ensembles differentially encode the commonalities and differences between situations. *J Neurosci* **31**, 12270-12281 (2011). https://doi.org:10.1523/JNEUROSCI.1671-11.2011

65 Markus, E. J. *et al.* Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *J Neurosci* **15**, 7079-7094 (1995).

66 Singer, A. C., Karlsson, M. P., Nathe, A. R., Carr, M. F. & Frank, L. M. Experience-dependent development of coordinated hippocampal spatial activity representing the similarity of related locations. *J Neurosci* **30**, 11586-11604 (2010). https://doi.org:10.1523/JNEUROSCI.0926-10.2010

67 McKenzie, S. *et al.* Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron* **83**, 202-215 (2014). https://doi.org:10.1016/j.neuron.2014.05.019

68 O'Keefe, J. & Krupic, J. Do hippocampal pyramidal cells respond to nonspatial stimuli? *Physiol Rev* **101**, 1427-1456 (2021). https://doi.org:10.1152/physrev.00014.2020

69 Knierim, J. J. From the GPS to HM: Place cells, grid cells, and memory. *Hippocampus* **25**, 719-725 (2015). https://doi.org:10.1002/hipo.22453

70 Moita, M. A., Rosis, S., Zhou, Y., LeDoux, J. E. & Blair, H. T. Hippocampal place cells acquire location-specific responses to the conditioned stimulus during auditory fear conditioning. *Neuron* **37**, 485-497 (2003). https://doi.org:10.1016/s0896-6273(03)00033-3

71 Shan, K. Q., Lubenov, E. V., Papadopoulou, M. & Siapas, A. G. Spatial tuning and brain state account for dorsal hippocampal CA1 activity in a non-spatial learning task. *eLife* **5** (2016). https://doi.org:10.7554/eLife.14321

72 Wood, E. R., Dudchenko, P. A. & Eichenbaum, H. The global record of memory in hippocampal neuronal activity. *Nature* **397**, 613-616 (1999). https://doi.org:10.1038/17605

73 Kennedy, P. J. & Shapiro, M. L. Motivational states activate distinct hippocampal representations to guide goal-directed behaviors. *Proc Natl Acad Sci U S A* **106**, 10805-10810 (2009). https://doi.org:10.1073/pnas.0903259106

74 Holscher, C., Jacob, W. & Mallot, H. A. Reward modulates neuronal activity in the hippocampus of the rat. *Behav Brain Res* **142**, 181-191 (2003).

75 Nieh, E. H. *et al.* Geometry of abstract learned knowledge in the hippocampus. *Nature* **595**, 80-84 (2021).

76 Johnston, W. J., Fine, J. M., Yoo, S. B. M., Ebitz, R. B. & Hayden, B. Y. Semi-orthogonal subspaces for value mediate a binding and generalization trade-off. *Nature Neuroscience*, 1-13 (2024).

77 Machens, C. K., Romo, R. & Brody, C. D. Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *Journal of Neuroscience* **30**, 350-360 (2010).

78 Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature* **503**, 78-84 (2013).

79 Flesch, T., Juechems, K., Dumbalska, T., Saxe, A. & Summerfield, C. Orthogonal representations for robust context-dependent task performance in brains and neural networks. *Neuron* **110**, 1258-1270. e1211 (2022).

80 Mizumori, S. J., Canfield, J. G. & Yeshenko, O. Parallel and interrelated neural systems underlying adaptive navigation. *Integrative and Comparative Biology* **45**, 547-554 (2005).

81 Mizumori, S. J., Cooper, B. G., Leutgeb, S. & Pratt, W. E. A neural systems analysis of adaptive navigation. *Mol Neurobiol* **21**, 57-82 (2000). https://doi.org:10.1385/MN:21:1-2:057

82 Mizumori, S. & Smith, D. Directing Neural Representations of Space. *Animal Spatial Cognition: Comparative, Neural, and Computational Approaches* (2006).

83 Buzsáki, G. & Moser, E. I. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature neuroscience* **16**, 130-138 (2013).

84 Bunsey, M. & Eichenbaum, H. Conservation of hippocampal memory function in rats and humans. *Nature* **379**, 255-257 (1996). https://doi.org:10.1038/379255a0

85 Safaie, M. *et al.* Preserved neural dynamics across animals performing similar behaviour. *Nature* **623**, 765-771 (2023). https://doi.org:10.1038/s41586-023-06714-0

86 Gallego, J. A. *et al.* Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature communications* **9**, 1-13 (2018).

87 Talpir, I. & Livneh, Y. Stereotyped goal-directed manifold dynamics in the insular cortex. *Cell reports* **43** (2024).

88 Kogan, E., Lu, J. & Zuo, Y. Cortical circuit dynamics underlying motor skill learning: from rodents to humans. *Frontiers in Molecular Neuroscience* **16**, 1292685 (2023).

89 Manoli, D. S., Meissner, G. W. & Baker, B. S. Blueprints for behavior: genetic specification of neural circuitry for innate behaviors. *Trends in neurosciences* **29**, 444-451 (2006).

90 LeDoux, J. & Daw, N. D. Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nature Reviews Neuroscience* **19**, 269-282 (2018).

91 LeDoux, J. E. As soon as there was life, there was danger: the deep history of survival behaviours and the shallower history of consciousness. *Philosophical Transactions of the Royal Society B* **377**, 20210292 (2022).

92 Harris, K. D. & Shepherd, G. M. The neocortical circuit: themes and variations. *Nature neuroscience* **18**, 170-181 (2015).

93 Shmuelof, L. & Krakauer, J. W. Are we ready for a natural history of motor learning? *Neuron* **72**, 469-476 (2011).

94 Tierney, A. Evolutionary implications of neural circuit structure and function. *Behavioural processes* **35**, 173-182 (1995).

95 Martínez-García, F. & Lanuza, E. Evolution of vertebrate survival circuits. *Current Opinion in Behavioral Sciences* **24**, 113-123 (2018).

96 Sweis, B. M., Mau, W., Rabinowitz, S. & Cai, D. J. Dynamic and heterogeneous neural ensembles contribute to a memory engram. *Curr Opin Neurobiol* **67**, 199-206 (2021). https://doi.org:10.1016/j.conb.2020.11.017

97 Davidson, T. J., Kloosterman, F. & Wilson, M. A. Hippocampal replay of extended experience. *Neuron* **63**, 497-507 (2009). https://doi.org:10.1016/j.neuron.2009.07.027 S0896-6273(09)00582-0 [pii]

98 Rubin, R. D., Watson, P. D., Duff, M. C. & Cohen, N. J. The role of the hippocampus in flexible cognition and social behavior. *Frontiers in human neuroscience* **8**, 742 (2014).

99 Poulter, S., Hartley, T. & Lever, C. The neurobiology of mammalian navigation. *Current Biology* **28**, R1023-R1042 (2018).

100 Tucker, M. A., Ord, T. J. & Rogers, T. L. Evolutionary predictors of mammalian home range size: body mass, diet and the environment. *Global Ecology and Biogeography* **23**, 1105-1114 (2014).

101 Schilder, B. M., Petry, H. M. & Hof, P. R. Evolutionary shifts dramatically reorganized the human hippocampal complex. *Journal of Comparative Neurology* **528**, 3143-3170 (2020).

102 Sherry, D. F., Jacobs, L. F. & Gaulin, S. J. Spatial memory and adaptive specialization of the hippocampus. *Trends in neurosciences* **15**, 298-303 (1992).

103 Pravosudov, V. V. & Roth II, T. C. Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. *Annual Review of Ecology, Evolution, and Systematics* **44**, 173-193 (2013).

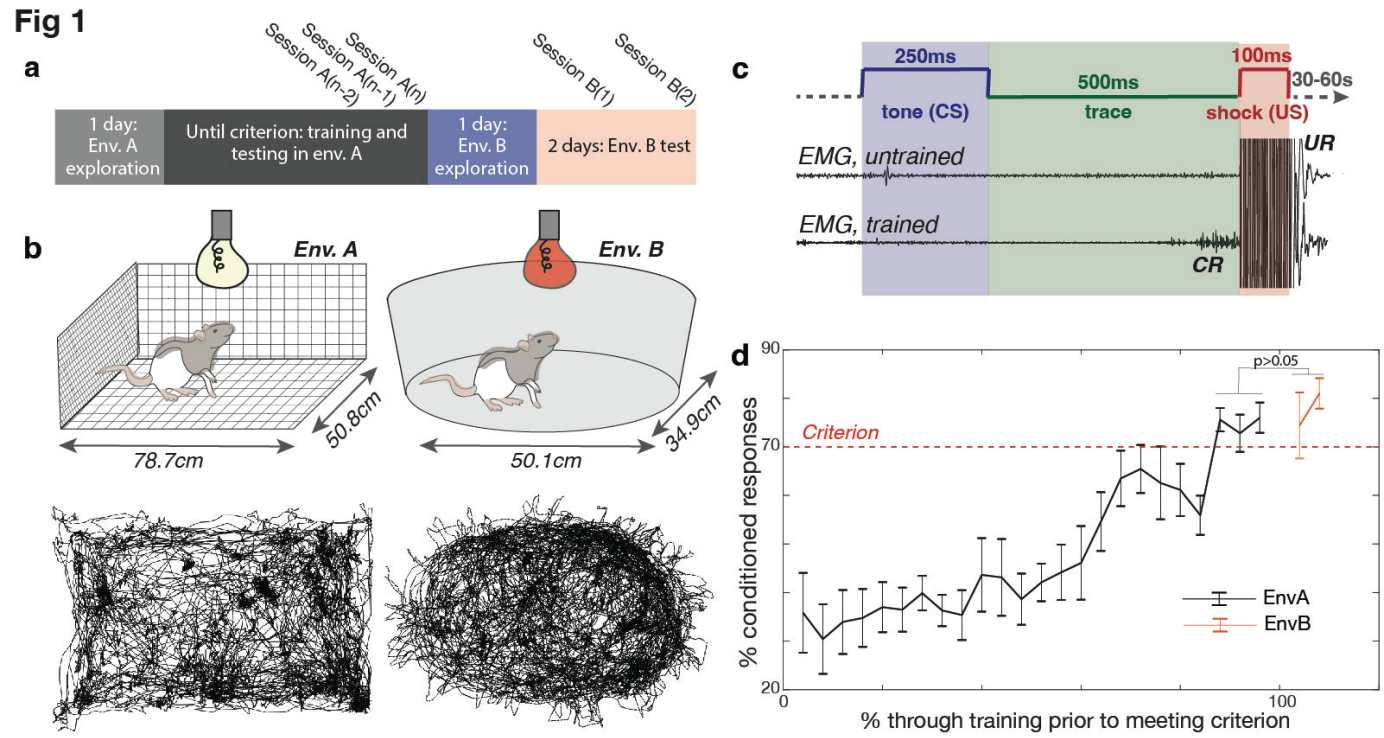


Fig 1. Experimental training paradigm and results.

a. Animals explored environment A for one session before undergoing eyeblink conditioning until they reached the learning criterion. Criterion was defined as achieving 70% conditioned responses (CRs) in 50 trials across three consecutive training sessions (A(n-2), A(n-1), and A(n)) or averaging over 70% CRs across the previous four training sessions. After meeting the criterion in environment A, animals were allowed one session of exploration in environment B, followed by two test sessions of trace eyeblink conditioning (B(1) and B(2)) in environment B.

b. (Top) Schematics of environments A and B. Environment A is a rectangular enclosure with wire walls, floor, and ceiling, lit with white light, and unscented. Environment B is oval-shaped with solid white floors and walls, without a ceiling, lit with red light, and scented with clove oil. Both environments provided distal cues visible from the top and

759 sides. (Bottom) Animal trajectories in environments A and B during a single session,
760 extracted using DeepLabCut.

761 **c.** Trace eyeblink conditioning (tEBC) paradigm. A 250 ms tone (conditioned stimulus, CS)
762 was followed by a 500 ms trace interval, then a 100 ms eyelid shock (unconditioned
763 stimulus, US). Eyelid activity was recorded using an EMG electrode implanted above the
764 eye. Untrained animals only blinked in response to the US (unconditioned response, UR);
765 trained animals began blinking during the trace interval after the CS and before the US
766 (conditioned response, CR).

767 **d.** Performance of animals (n=5) in the tEBC task. Animals learned tEBC while freely
768 moving in environment A and successfully transferred this learning to environment B.
769 The dotted line indicates the performance criterion. No significant difference was found
770 between performance in the criterion sessions in environment A (mean 74.75 ± 6.49%)
771 and the test sessions in environment B (mean 77.70 ± 11.68%; two-tailed t-test, $t(24) = -$
772 0.83, $p > 0.05$). Error bars represent standard error.

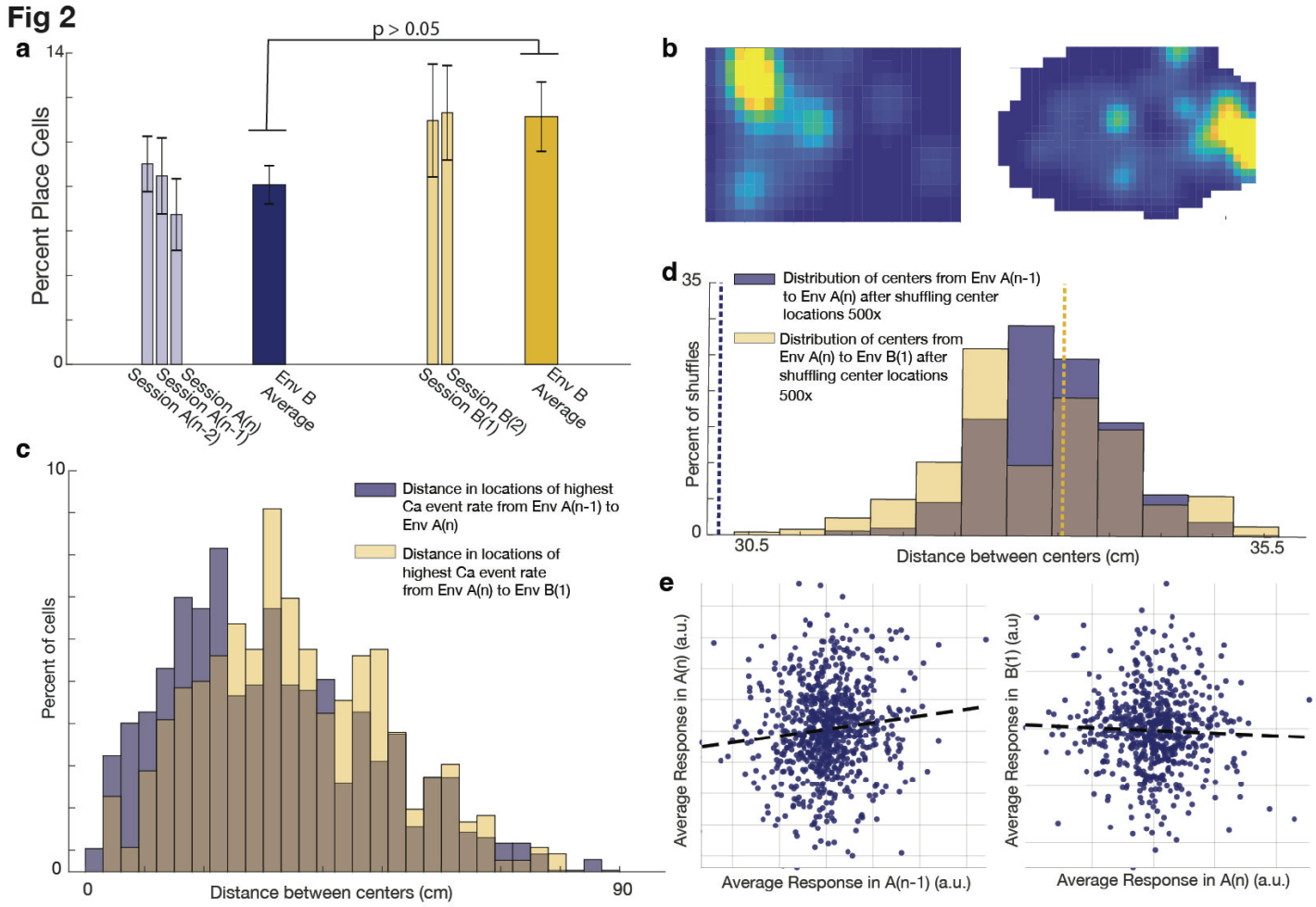


Fig 2. Place cells remap between environment A and environment B.

a. Percent of place cells during criterion sessions in environment A and test sessions in environment B. Light-colored bars represent averages within individual sessions, and dark-colored bars represent overall averages in environments A and B. Overlaid bars indicate standard deviation. On average, 9.3% ± 4.2% of cells were classified as place cells. There was no significant difference in the percentage of place cells between environments A and B (average in environment A: 8.1% ± 3.3%; average in environment B: 11.1% ± 4.9%; two-tailed t-test, $t(23) = -1.9$, $p > 0.05$).

782 **b.** Example activity map of a specific single cell in environments A and B. Yellow indicates
783 the highest firing rates. This cell exhibited remapping between environments, showing
784 different place fields relative to external cues in the two contexts.
785 **c.** Distribution of distances between place field centers (determined by the highest calcium
786 event rate), comparing sessions A(n) and A(n-1) versus sessions A(n) and B(1). Place
787 field centers shifted significantly more when the animal was moved to environment B
788 compared to shifts between two consecutive sessions in environment A (Wilcoxon rank
789 sum test, $p = 0.002$; two-sided t-test, $t(1430) = -2.5$, $p = 0.01$; two-sample Kolmogorov-
790 Smirnov test, $p = 3.6 \times 10$). The brown shaded area represents the overlap between
791 distance histograms.
792 **d.** Distribution of the median distance between place field centers after shuffling center
793 locations 500 times. The actual median value for session A(n) to A(n-1) was smaller than
794 all shuffled medians ($p = 0$, dashed blue line), while the median for session A(n) to B(1)
795 was greater than or equal to 56% of shuffled medians ($p = 0.56$, dashed yellow line). The
796 brown shaded area represents the overlap between shuffled distributions.
797 **e.** Population vector correlation (PVC) based on calcium events for sessions A(n-1), A(n),
798 and B(1). A significant positive correlation was found between sessions A(n-1) and A(n)
799 for cells present in both sessions ($p = 0.0023$, $r = 0.11$). In contrast, there was no
800 significant correlation between sessions A(n) and B(1) for shared cells ($p > 0.05$, $r = -$
801 0.04). These findings suggest that calcium event patterns are significantly similar
802 between sessions A(n-1) and A(n) but not between session A(n) and B(1). Dashed lines
803 represent lines of best fit. (a.u. = arbitrary units).
804

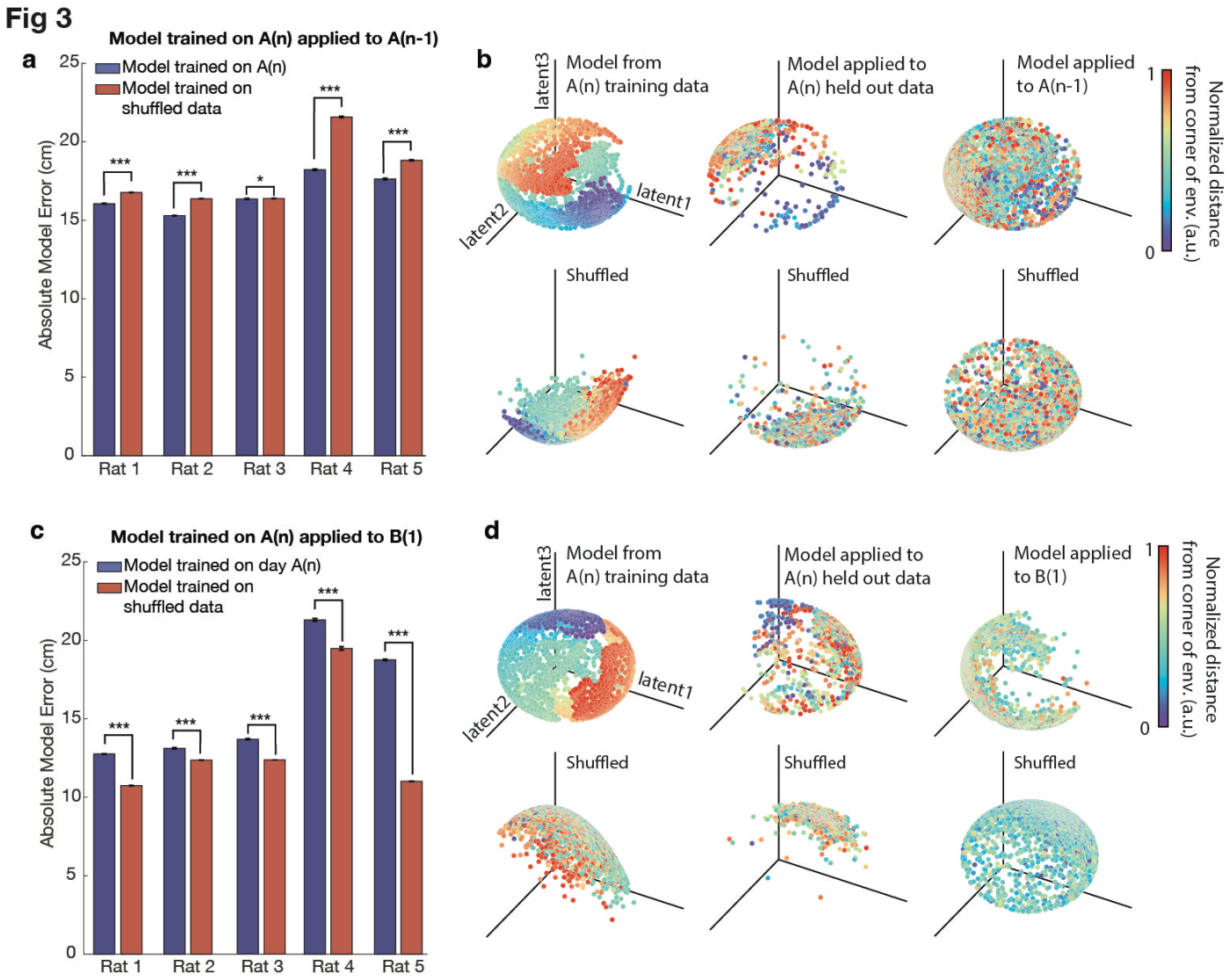


Fig. 3. A model trained in environment A can decode positions within environment A but not in environment B.

a. A model trained on calcium trace and position data from session A(n), using cells present in both A(n) and A(n-1), predicted positions in environment A(n-1) with significantly greater accuracy than that of a model trained on shuffled position data. The model was run 500 times, and the accuracy of predictions was assessed. The model trained on actual data significantly outperformed the shuffled model across rats (double-sided t-tests: Rat 1: $t(998) = -34.3$, $p = 5.0 \times 10^{-171}$; Rat 2: $t(998) = -72.5$, $p = 0$; Rat 3: $t(998) = -1.96$, $p =$

0.05; Rat 4: $t(998) = -53.7$, $p = 2.6 \times 10^{-299}$; Rat 5: $t(998) = -20.0$, $p = 4.1 \times 10^{-75}$). Error bars represent standard error. A single asterisk (*) indicates $p \leq 0.05$, and three asterisks (***) indicate $p < 10^{-74}$.

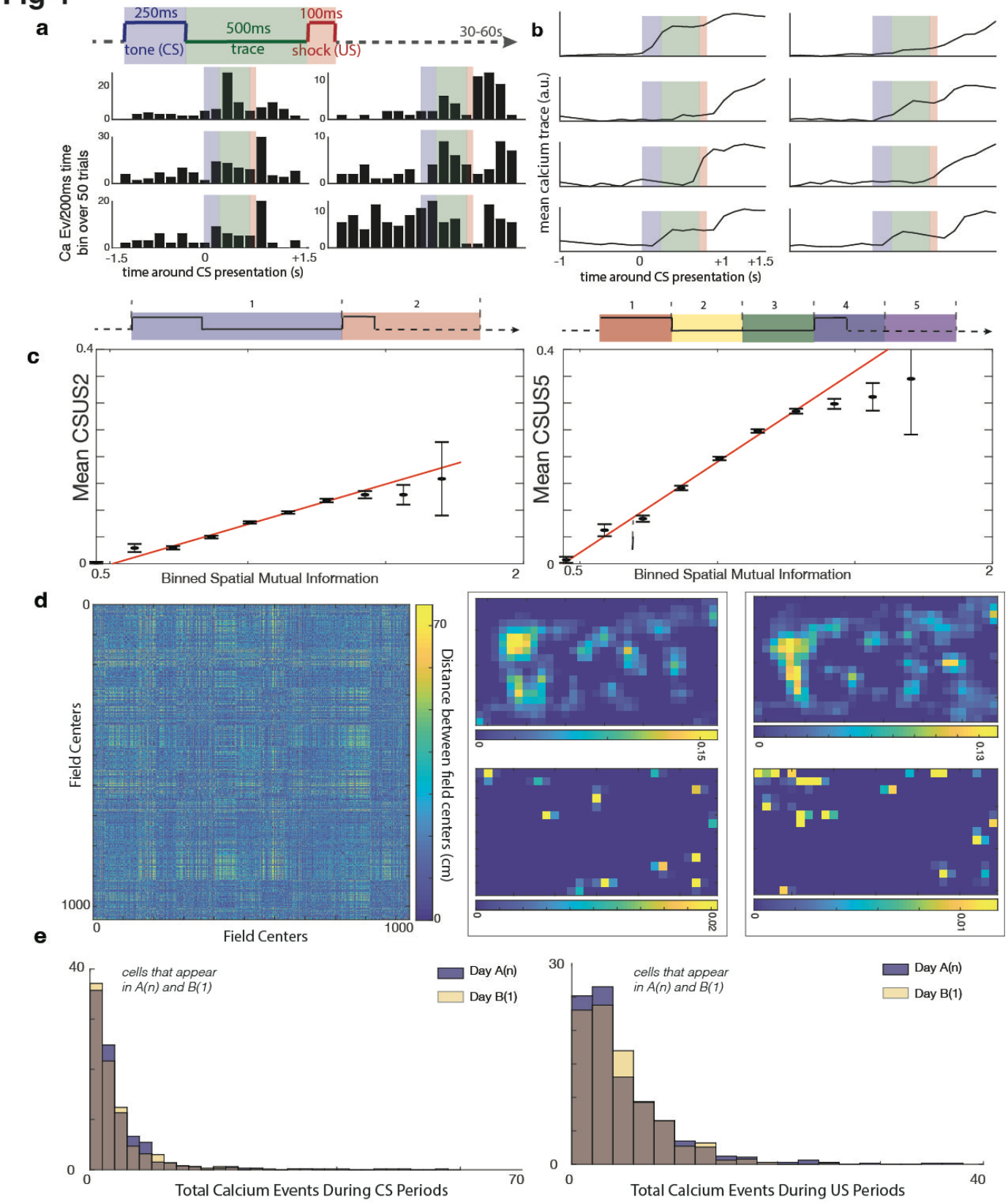
b. (Top row) Visualization of model performance for decoding animal position in environment A(n). Left: The trained model demonstrated on the training data from session A(n). Middle: The same model applied to held-out trace data (25%) from session A(n). Right: The model applied to predict animal position in session A(n-1). (Bottom row) The same models trained on shuffled position data. Shown here is the model for Rat 4, where the model trained on real data significantly outperformed model trained on shuffled data for decoding position in session A(n-1) (500 simulations, $t(998) = -53.7$, $p = 2.6 \times 10^{-299}$). For visualization purposes, distance from the corner of the environment is plotted using normalized values in arbitrary units [a.u.].

c. A model trained on data from session A(n), using cells present in both A(n) and B(1), was applied to environment B(1). The model's predictions were significantly less accurate than those of a model trained on shuffled position data (double-sided t-tests: Rat 1: $t(998) = 84.1$, $p = 0$; Rat 2: $t(998) = 18.8$, $p = 7.4 \times 10^{-68}$; Rat 3: $t(998) = 74.7$, $p = 0$; Rat 4: $t(998) = 13.1$, $p = 2.0 \times 10^{-36}$; Rat 5: $t(998) = 154.4$, $p = 0$). Error bars represent standard error. Three asterisks (***) indicate $p < 10^{-35}$.

d. (Top row) Model trained on position and calcium trace data from session A(n), using cells present in both A(n) and B(1). Left: The model demonstrated on the training data from session A(n). Middle: The same model applied to held-out trace data (25%) from session A(n). Right: The model applied to predict the animal position in session B(1). (Bottom row) The same models trained on shuffled position data. Shown here is the model for Rat 4, where the shuffled model performed significantly better than the model

trained on session A(n) for decoding position in session B(1) (500 simulations, $t(998) = 13.1$, $p = 2.0 \times 10^{-36}$). For visualization purposes, distance from the corner of the environment is plotted using normalized values in arbitrary units [a.u.].

Fig 4


a
250ms
tone (CS)
500ms
trace
100ms
shock (US)
30-60s
Ca Ev/200ms time bin over 50 trials
time around CS presentation (s)
b
mean calcium trace (a.u.)
time around CS presentation (s)
c
Mean CSUS2
Binned Spatial Mutual Information
Mean CSUS5
Binned Spatial Mutual Information
d
Field Centers
Field Centers
Distance between field centers (cm)
e
cells that appear in A(n) and B(1)
Day A(n)
Day B(1)
Total Calcium Events During CS Periods
cells that appear in A(n) and B(1)
Day A(n)
Day B(1)
Total Calcium Events During US Periods

844 **Figure 4. Identifying CS and US modulation in individual cells**

845 **a.** Peristimulus time histograms (PSTHs) of six example cells responding to the CS, US,
846 and/or trace period. Bars represent calcium events summed across 50 trials. The
847 shaded regions indicate different stimulus periods: blue (CS), green (trace), and red
848 (US).

849 **b.** Mean calcium traces of eight example cells in response to the CS, US, and/or trace
850 period, averaged across 50 trials. Shaded regions are consistent with panel a (a.u. =
851 arbitrary units).

852 **c.** Scatter plots illustrating the significant positive correlation between spatial mutual
853 information (MI) and CSUS mutual information (MI). Data points represent binned
854 observations with error bars showing standard errors. The solid red line represents the
855 best fit from a linear regression analysis of the original unbinned dataset. Left: CSUS-
856 MI2 ($r^2 = 0.04$, $p = 2.1 \times 10^{-104}$). Right: CSUS-MI5 ($r^2 = 0.09$, $p = 1.7 \times 10^{-237}$).

857 **d.** Mosaic plots showing the odds ratios for cells modulated by spatial and/or CSUS
858 information. Cells were classified as significantly modulated if their mutual
859 information exceeded 95% of shuffled values. Left: Cells with significant spatial
860 modulation had a 1.35 times higher chance of being significantly modulated by
861 CSUS-MI2 compared to cells without spatial modulation (Fisher's exact test, p =
862 0.001). Right: Similar results for CSUS-MI5, where cells with significant spatial
863 modulation had a 1.27 times higher chance of being significantly modulated by
864 CSUS-MI5 (Fisher's exact test, p = 0.002).

865 **e.** Place field remapping during conditioning and non-conditioning periods. (Left) Heat
866 map comparing the spatial distances between place field centers during conditioning
867 (e.g., during CS/US trials) versus non-conditioning periods (e.g., intertrial intervals).

Color represents the absolute difference in distances between field centers: regions with minimal color variation (blue) suggest similar place field centers between the two conditions, while more yellow areas represent significant differences in distance (Mantel statistic = 534.58, $p > 0.05$). (Right) Example of two cells showing spatial calcium activity. Top: Heat maps showing calcium event rates during periods of movement but not conditioning (e.g., intertrial intervals). Bottom: Calcium event rates during CS/US conditioning periods. Dark blue signifies the lowest event rate, and yellow represents the highest event rate. These results show that calcium events during trials are not confined to a cell's place field.

f. Distribution of calcium events during CS (left) and US (right) periods in sessions A(n) (blue) and B(1) (yellow). No significant difference was observed in hippocampal firing during the CS (left) or US (right) periods between environments A and B (CS: two-tailed t-test, $t(1174) = 0.68$, $p > 0.05$; US: two-tailed t-test, $t(1174) = 1.20$, $p > 0.05$; KS tests, $p > 0.05$ for both comparisons).

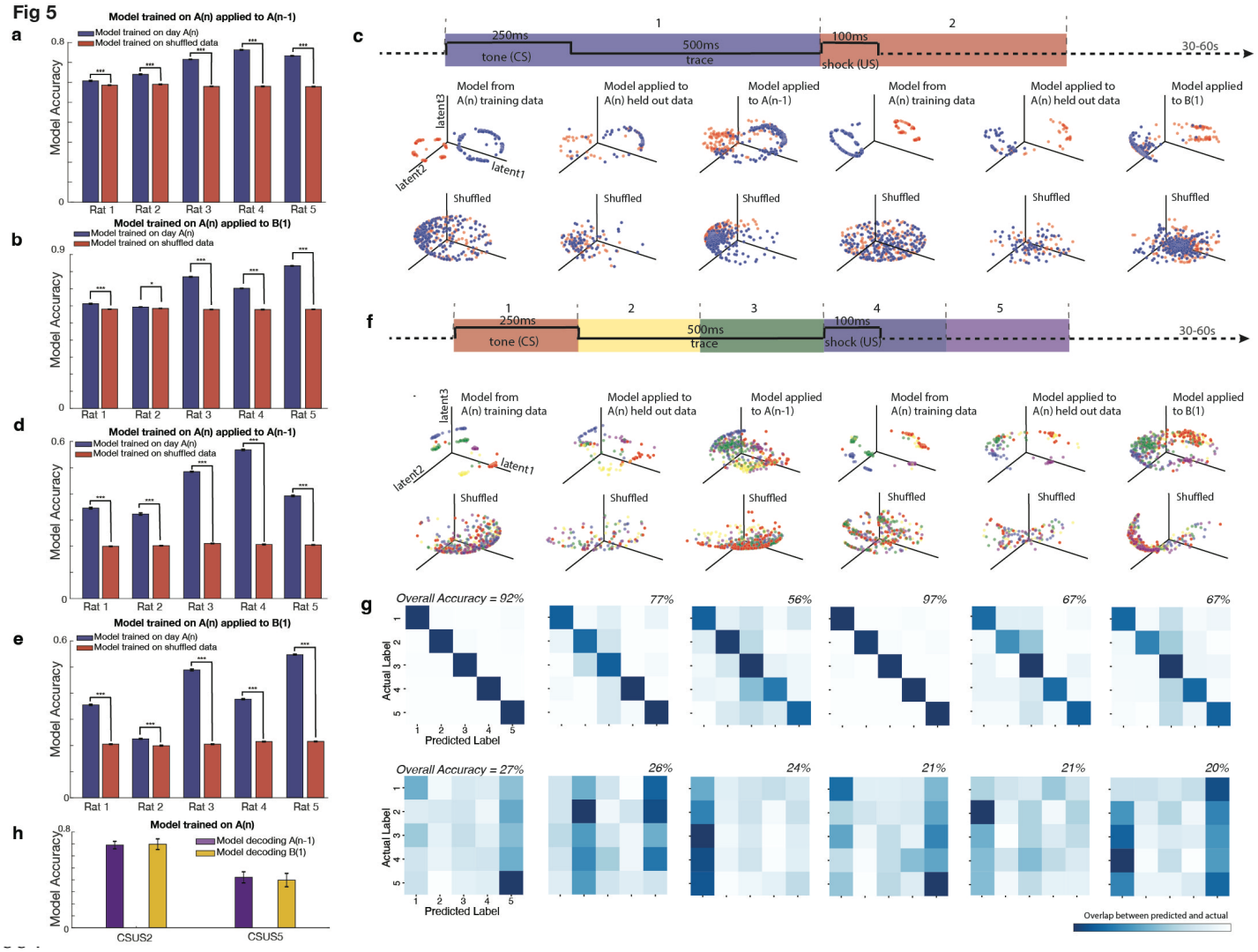


Fig. 5. A model trained in environment A can decode CS and US periods in both environment A and B at above chance levels, including fine-grained temporal decoding.

a. A CEBRA model was trained using calcium imaging data and time-stamped CS/US periods from environment A, using only cells that were recorded in both environments A and B. The model was used to decode whether the animal was in a CS or US period in another session in environment A; all models successfully decoded these periods compared to shuffled data (all double-sided t-tests: Rat 1: $t(998) = 6.7$, $p = 2.5 \times 10^{-11}$; Rat 2: $t(998) = 16.1$, $p = 7.4 \times 10^{-52}$; Rat 3: $t(998) = 83.5$, $p = 0$; Rat 4: $t(998) = 80.1$, $p =$

894 0; Rat 5: $t(998) = 61.0$, $p = 0$). Bars indicate standard error; significance denoted as ***p
895 $< 10^{-10}$.

896 **b.** The same models from panel a were now applied to environment B; they significantly
897 outperformed chance level (all double-sided t-tests: Rat 1: $t(998) = 10.4$, $p = 2.6 \times 10^{-24}$;
898 Rat 2: $t(998) = 2.7$, $p = 7.6 \times 10^{-3}$; Rat 3: $t(998) = 75.3$, $p = 0$; Rat 4: $t(998) = 63.7$, $p = 0$;
899 Rat 5: $t(998) = 106.3$, $p = 0$). Error bars represent standard error, *$p < 10^{-3}$, ***$p < 10^{-10}$.

900 **c.** (Top row) A CEBRA model trained on CS/US periods, divided into two time bins (data
901 from Rat 4). For the top first three graphs, the model was trained using position and
902 calcium trace data from session A(n), using cells present in both A(n) and A(n-1). First
903 graph: The trained model applied to training data from session A(n). Second graph: The
904 model applied to decode CS/US periods from held-out data (25%) from session A(n).
905 Third graph: The model applied to session A(n-1). For the top last three graphs, the
906 model was trained using position and calcium trace data from session A(n), using cells
907 present in both A(n) and B(1). Fourth graph: The trained model applied to training data
908 from session A(n). Fifth graph: The model applied to decode CS/US periods from held-
909 out data (25%) from session A(n). Sixth graph: The model applied to session B(1).
910 (Bottom row) The same model trained on shuffled data. The model trained on data from
911 session A(n) significantly outperformed the model trained on shuffled data in decoding
912 both session A(n-1) and session B(1) (double-sided t-tests, $t(998) = 83.5$, $p = 0$, and
913 $t(998) = 75.3$, $p = 0$, respectively).

914 **d.** A CEBRA model was trained on data from environment A to decode temporal order
915 within the CS, trace, and US periods split into five divisions, applied to an alternate
916 session in environment A (session A(n-1)). The model significantly outperformed chance
917 (all double-sided t-tests for accuracy: Rat 1: $t(998) = 41.5$, $p = 3.3 \times 10^{-220}$; Rat 2: $t(998)$

= 28.7, $p = 7.8 \times 10^{-133}$; Rat 3: t(998) = 122.6, p = 0; Rat 4: t(998) = 118.5, p = 0; Rat 5: t(998) = 62.6, p = 0). Bars indicate standard error; significance denoted as ***$p < 10^{-10}$.

e. The same five models from panel d were now applied to environment B and outperformed models trained on shuffled data in decoding the temporal aspects of the CS/US periods, indicating that fine-grained temporal encoding is stable across environments (all double-sided t-tests for accuracy: Rat 1: t(998) = 55.1, $p = 4.9 \times 10^{-305}$; Rat 2: t(998) = 9.3, $p = 1.1 \times 10^{-19}$; Rat 3: t(998) = 71.4, p = 0; Rat 4: t(998) = 62.6, p = 0; Rat 5: t(998) = 106.2, p = 0). Bars indicate standard error; significance denoted as ***$p < 10^{-10}$.

f. Same analysis as in panel c, but for a CEBRA model trained on CS/US periods split into five time bins. Top: The five divisions of the conditioning period are shown. Data from Rat 3 show that the model trained on session A(n) outperformed models trained on shuffled data when decoding both session A(n-1) and session B(1) (double-sided t-tests, t(998) = 88.9, p = 0, and t(998) = 71.4, p = 0, respectively).

g. Confusion matrices displaying CEBRA decoding of five CS/US time bins, as shown in panel f. Top row: Models trained on data from session A(n). Bottom row: Models trained on shuffled data. Darker colors along the diagonal indicate higher model accuracy.

h. Model accuracy for decoding sessions A(n-1) and B(1) using a model trained on data from session A(n) with CSUS2 and CSUS5 divisions. For both CSUS2 and CSUS5, a model trained in session A(n) decoded session B(1) with accuracy similar to that achieved when decoding session A(n-1) (CSUS2: double-sided t-test, t(8) = -0.13, $p > 0.05$; CSUS5: t(8) = 0.32, $p > 0.05$). Bars represent standard error

Fig 6

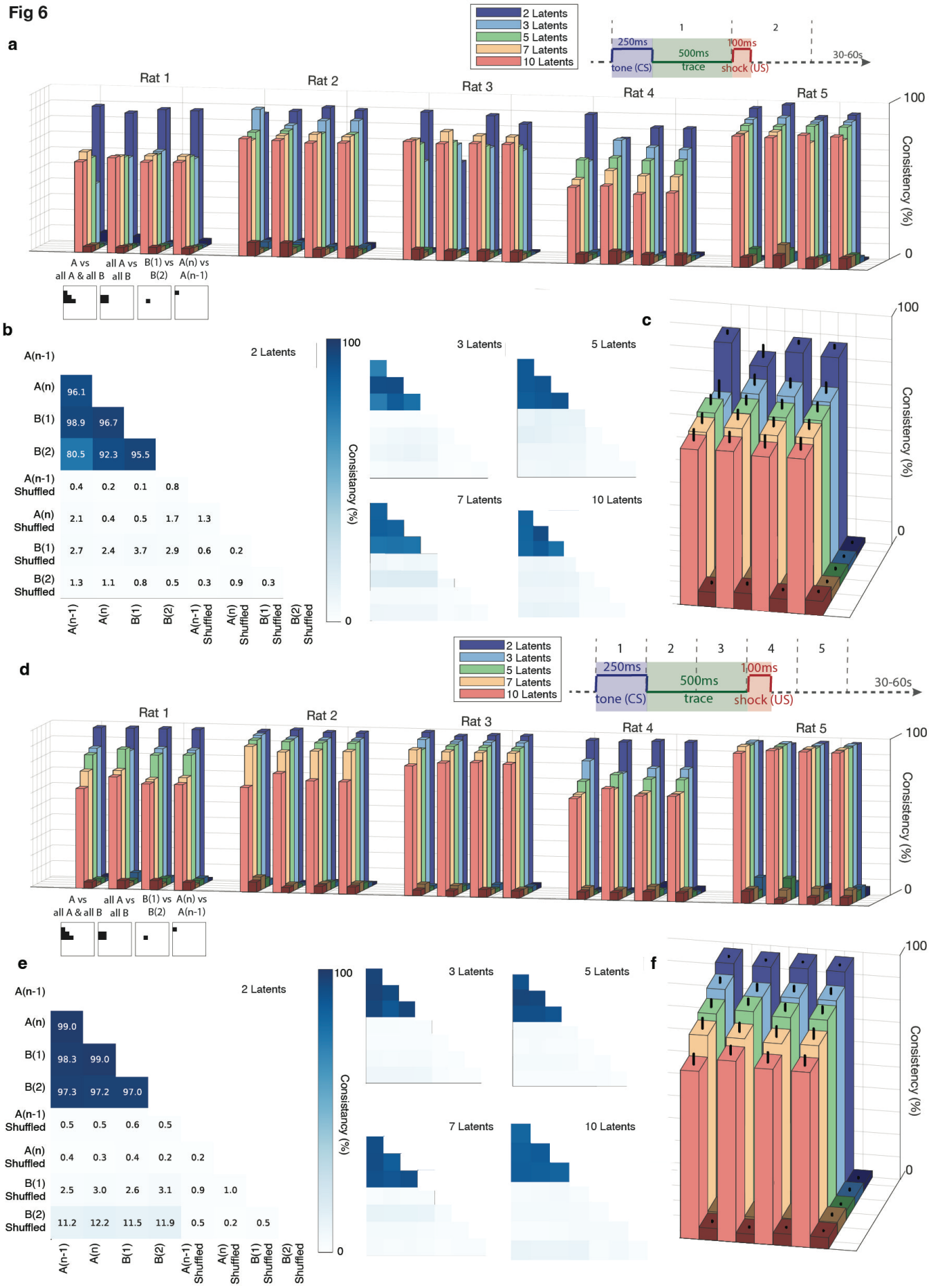

a
2 Latents
3 Latents
5 Latents
7 Latents
10 Latents
250ms
500ms
100ms
tone (CS)
trace
shock (US)
30-60s
Rat 1
Rat 2
Rat 3
Rat 4
Rat 5
Consistency (%)
100
0
A vs all A & all B
all A vs all B
B(1) vs B(2)
A(n) vs A(n-1)
b
A(n-1)
A(n) 96.1
B(1) 98.9 96.7
B(2) 80.5 92.3 95.5
A(n-1) Shuffled 0.4 0.2 0.1 0.8
A(n) Shuffled 2.1 0.4 0.5 1.7 1.3
B(1) Shuffled 2.7 2.4 3.7 2.9 0.6 0.2
B(2) Shuffled 1.3 1.1 0.8 0.5 0.3 0.9 0.3
A(n-1) A(n) B(1) B(2) A(n-1) Shuffled A(n) Shuffled B(1) Shuffled B(2) Shuffled
2 Latents
3 Latents
5 Latents
7 Latents
10 Latents
Consistency (%)
100
0
c
d
e
A(n-1)
A(n) 99.0
B(1) 98.3 99.0
B(2) 97.3 97.2 97.0
A(n-1) Shuffled 0.5 0.5 0.6 0.5
A(n) Shuffled 0.4 0.3 0.4 0.2 0.2
B(1) Shuffled 2.5 3.0 2.6 3.1 0.9 1.0
B(2) Shuffled 11.2 12.2 11.5 11.9 0.5 0.2 0.5
f

Fig 6. High consistency in neural representations between environments A and B.

a. Consistency scores for each rat, calculated with 2, 3, 5, 7, and 10 latents. Lighter bars represent consistency percentage for actual data, while adjacent darker bars represent consistency for shuffled data. The schematic below the x-axis illustrates the data sets included in each comparison (for position and labelling of data sets see figure 6b). For the CSUS2 division of conditioning periods (split into CS and US components), all five rats show significantly higher consistency between environments A and B in the actual data compared to shuffled data. This consistency remains significant with up to 10 latents.

b. Example consistency measurements from Rat 5 for CSUS2 with 2, 3, 5, 7, and 10 latents.

c. Average consistency scores across all rats for CSUS2. The figure legend follows the format of panel a. Bars represent standard error. Consistency scores for actual data versus shuffled data were significantly different across all latent dimensions (all double-sided t-tests: 2 latents, $p < 1 \times 10^{-5}$; 3 latents, $p < 1 \times 10^{-4}$; 5 latents, $p < 1 \times 10^{-5}$; 7 latents, $p < 1 \times 10^{-5}$; 10 latents, $p < 1 \times 10^{-4}$).

d. Consistency scores for each rat, as in panel a, but for CSUS5, where the conditioning period is divided into five time bins. All five rats show significantly higher consistency between environments A and B in the actual data compared to shuffled data; this is maintained across latents for up to 10.

e. Example consistency measurements from Rat 3 for CSUS5 with 2, 3, 5, 7, and 10 latents.

f. Average consistency scores across all rats for CSUS5. The figure legend matches that of panel c. Bars represent standard error. Significant differences between actual and

shuffled data were observed for all latent dimensions (all double-sided t-tests: 2 latents, $p < 1 \times 10^{-12}$; 3 latents, $p < 1 \times 10^{-7}$; 5 latents, $p < 1 \times 10^{-6}$; 7 latents, $p < 1 \times 10^{-5}$; 10 latents, $p < 1 \times 10^{-5}$).

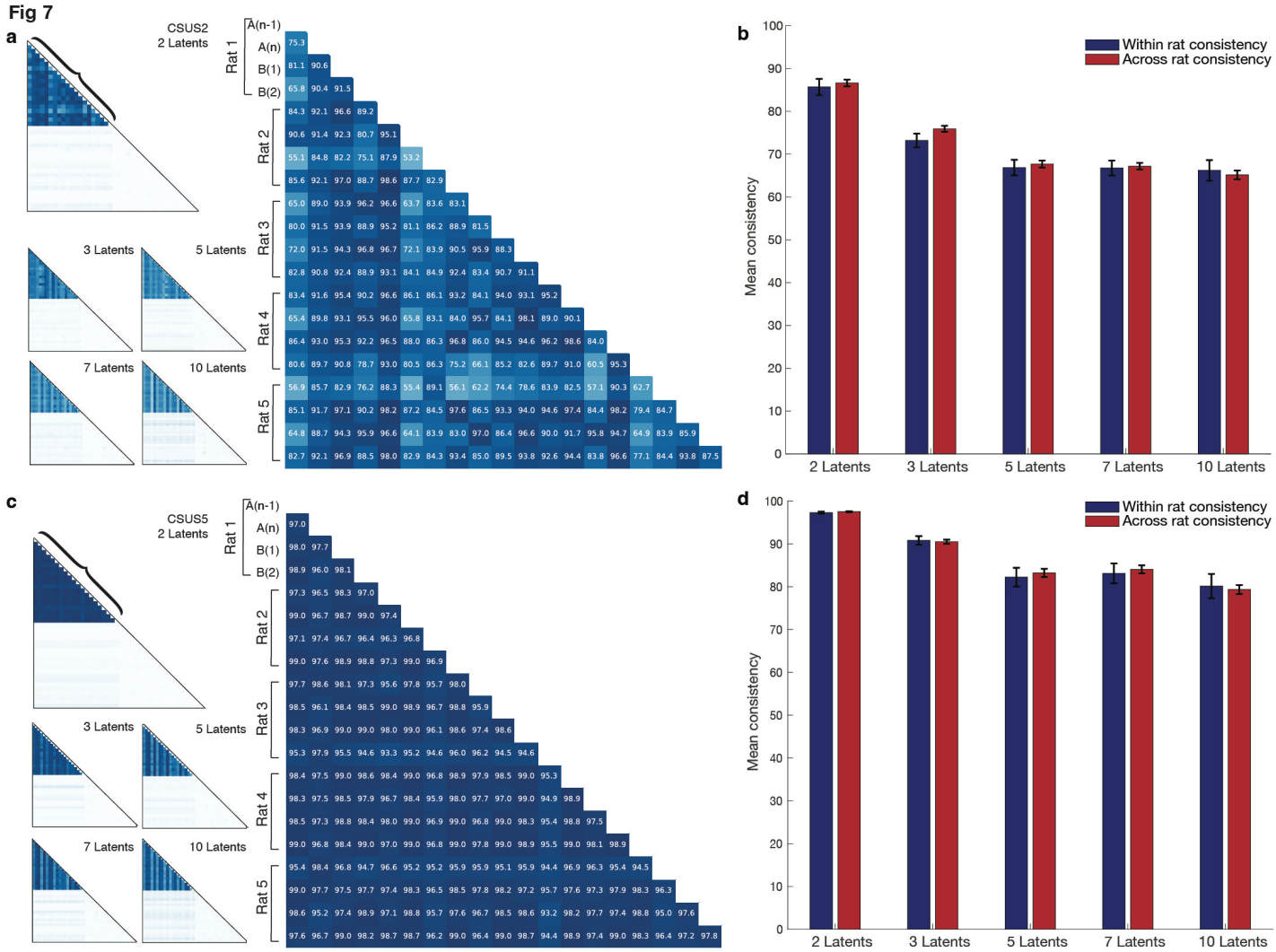


Figure 7. High degree of consistency between the conditioning representations across animals

a. Consistency across animals for CSUS2 with 2, 3, 5, 7, and 10 latents. The larger graph highlights the bracketed area from the graph with two latents.

b. Conditioning period is divided into 2 time bins (CSUS2): the similarity between models across animals is not significantly different from the similarity between models for an individual animal, regardless of the number of latents (2 latents: $t(188) = -0.45$, $p > 0.05$;

3 latents: $t(188) = -1.57$, $p > 0.05$; 5 latents: $t(188) = -0.40$, $p > 0.05$; 7 latents: $t(188) = -0.22$, $p > 0.05$; 10 latents: $t(188) = 0.40$, $p > 0.05$). Error bars represent standard error.

c. Consistency across animals for CSUS5 with 2, 3, 5, 7, and 10 latents. The larger graph highlights the bracketed area from the graph with two latents.

d. Conditioning period is divided into 5 time bins (CSUS5); the similarity across animal models remains comparable to that for within-animal models, regardless of the number of latents (2 latents: $t(188) = -0.79$, $p > 0.05$; 3 latents: $t(188) = 0.25$, $p > 0.05$; 5 latents: $t(188) = -0.42$, $p > 0.05$; 7 latents: $t(188) = 0.70$, $p > 0.05$; 10 latents: $t(188) = 0.30$, $p > 0.05$). Error bars represent standard error.

992 **Methods**

993 **LEAD CONTACT AND MATERIALS AVAILABILITY**

994 Questions and requests for information should be directed to and will be fulfilled by the
995 Lead Contact, Hannah Wirtshafter (hsw@northwestern.edu). This study did not generate new
996 unique reagents. The data that support the findings of this study are available from the
997 corresponding author.

998

999 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

1000 All procedures were performed within Northwestern Institutional Animal Care and Use
1001 Committee and NIH guidelines. Five male Long Evans rats (275–325 g) were sourced from
1002 Charles River Laboratories, injected with AAV9-GCaMP8m, implanted with a 2-mm GRIN lens,
1003 and trained and tested on eyeblink conditioning in two apparatuses (Fig. 1). Animals were
1004 individually housed in an animal facility with a 12/12 h light/dark cycle.

1005

1006 **METHOD DETAILS**

1007 **GCaMP7c injection, lens implantation, EMG implantation**

1008 GCaMP8 injection and lens implantation were completed as reported in Wirtshafter and
1009 Disterhoft, 2022 and Wirtshafter and Disterhoft, 2023[1,2]. Briefly, rats were anesthetized with
1010 isoflurane (induction 4%, maintenance 1-2%) and a craniotomy was performed at stereotaxic
1011 coordinates Bregma AP −4.00mm, ML 3.00mm. 0.06uL of GCaMP8m (obtained from
1012 AddGene, packaged AAV9 of pGP-AAV-syn-jGCaMP8m-WPRE, lot v175525, titer 1.3E+13
1013 GC/mL) was injected over 12 minutes (approximate coordinates Bregma AP −4.00mm, ML
1014 3mm, DV 2.95mm relative to skull); then the syringe was raised 0.2mm and an additional 0.6ul

1015 of GCaMP7 was injected. We repeated this process once more and at slightly different
1016 coordinates in the craniotomy hole, resulting in 4 total injections.
1017 We then aspirated tissue from the craniotomy site using a vacuum pump and 25 gauge
1018 needle. Tissue was aspirated up to and including the horizontal striations of the corpus collosum.
1019 A 2mm GRIN lens (obtained from Go!Foton, CLH lens, 2.00mm diameter, 0.448 pitch, working
1020 distance 0.30mm, 550nm wavelength) was then inserted into the craniotomy hole and cemented
1021 in place using dental acrylic. Animals were given buprenorphine (0.05mg/kg) and 20mL saline,
1022 taken off anesthesia, and allowed to recover in a clean cage placed upon a heat pad.
1023 Six to eight weeks after surgery, animals were again anesthetized with isoflurane and
1024 checked for GCaMP expression. If expression was seen, baseplates were attached using UV-
1025 curing epoxy and dental acrylic. Electrode implantation to record obicularis oculi
1026 electromyographic (EMG) activity occurred in the same surgery as baseplate attachment, as
1027 described previously[3,4]. Briefly, a connector containing 5 wires was cemented on the front of the
1028 animal's head: 4 wires were implanted directly above the eye in the surrounding muscle (2 for
1029 recording, 2 for electrical stimulation). An additional wire was attached to a connector attached
1030 to a ground screw located above the cerebellum; this screw was implanted during lens
1031 implantation surgery.
1032
1033 **Behavioral environment and training**
1034 Two behavioral apparatuses were used in these experiments: Environment A was a
1035 78.7cm x 50.8cm unscented rectangular enclosure with wire floor and walls and white lighting.
1036 Environment B was a 50.1cm x 34.9cm scented (with two dabs of clove essential oil on opposite
1037 walls) ovular enclosure with white solid floor and walls, and red lighting. Both environments
1038 were located at the same spot in the room relative to external cues (see Figures 1b and S1).

A tether containing a plug to relay the EMG activity and to deliver a shock to the rat's eye was attached to a the eyeblink connector on the rat's head. The miniscope was plugged into the cemented baseplate. The miniscope and EMG cords were all attached to a commutator for ease of animal movement.

The CS was a 250ms, 85dB free-field tone (5ms rise-fall time). The US was a 100ms shock directed to the left eye. Shock amount varied per session per animal and was calibrated, if needed, at the end of a training session for the next session's training. Shock level was deemed appropriate when a shock was met with a firm shake of the animal's head.

The trace interval was 500ms and the intertrial interval (ITI) was randomized between 30s and 60s, with a 45s average. EMG signal output was amplified (5000×) and filtered (100 Hz to 5 kHz), then digitized at 3 kHz and stored by computer.

A conditioned response (CR) was identified as an increase in integrated EMG activity that exceeded the baseline mean amplitude by more than four standard deviations, sustained for a minimum duration of 15ms. Baseline mean amplitude was calculated during the 500ms preceding CS onset. Additionally, the response had to commence at least 50ms after the conditioned stimulus (CS) onset and before the unconditioned stimulus (US) onset.

The animal's first exposure to each environment was a 38min exploration session, in which the animal was able to freely move and explore the environment without any conditioning (Figure 1a). Animals were then trained in one environment per session, with no more than one session per day, and were considered to have learned the task after reaching criterion (70% CRs in 50 trials) on three consecutive training sessions (termed 'criterion sessions') or when the previous four training sessions averaged over 70% (in this instance, only the final three of those sessions were considered 'criterion sessions'). Following the last session in environment A, the animal was given an exploratory session in environment B. The session after that, the animal was

1063 tested on eye blink conditioning in environment B, using the same parameters as used in
1064 environment A.
1065
1066 **Calcium imaging**
1067 Calcium imaging was completed as reported in Wirtshafter and Disterhoft, 2022 and
1068 Wirtshafter and Disterhoft, 2023[1,2]. Briefly, calcium imaging was done using UCLA V4
1069 Miniscopes[5,6], assembled with two 3mm diameter, 6mm FL achromat lens used in the objective
1070 module and one 4mm diameter, 10mm FL achromat lens used in the emission module.
1071
1072 **QUANTIFICATION AND STATISTICAL ANALYSIS**
1073 Means are presented as mean+-standard deviation. All analysis code is available
1074 at https://github.com/hsw28/ca_imaging and https://github.com/hsw28/Hannahs-CEBRAs. Code
1075 to create specific figures is also available at the former github repository.
1076
1077 **Position and speed analysis**
1078 Position was sampled by an overhead camera at 30Hz. Position tracking was done post-
1079 recording using DeepLabCut[7]. Position was then converted from pixels to cm. Position was
1080 smoothed using a Gaussian filter with standard deviation of 2cm. Speed was calculated by taking
1081 the hypotenuse of the coordinates one before and after the time of interest.
1082
1083 **Video pre-processing and cell identification**
1084 Video pre-processing and cell identification were performed as reported in Wirtshafter
1085 and Disterhoft, 2022 and Wirtshafter and Disterhoft, 2023[1,2]. In brief, videos were recorded with
1086 Miniscope software at 15frames/second. Video processing was done using CIATAH software[8].

1087 Videos were down sampled in space and normalized by subtracting the mean value of each
1088 frame from the frame. Each frame was then normalized using a bandpass FFT filter (70-
1089 100cycles/pixel) and motion corrected to a using TurboReg[9]. Videos were then converted to
1090 relative florescence (dF/F_0); F_0 was the mean over the entire video.

1091 Cells were automatically identified using CIATAH[8] using CNMF-E[10]. Images were
1092 filtered with a gaussian kernel of width 2 pixels and neuron diameter was set at a pixel size of 8.
1093 The threshold for merging neurons was set at a calcium trace correlation of 0.65; neurons were
1094 merged if their distances were smaller than 4 pixels and they had highly correlated spatial shapes
1095 (correlation>0.8) and small temporal correlations (correlation <0.4).

1096 In vivo calcium imaging involves detecting changes in intracellular calcium levels, which
1097 serve as proxies for neuronal activity. Calcium events refer to transient increases in calcium
1098 concentration above a threshold level; these crossings putatively correspond to spikes in neuronal
1099 firing. These events typically appear as peaks in the data and indicate an active response from the
1100 neuron. Calcium traces are continuous recordings of calcium levels over time. Thus, calcium
1101 events highlight specific neuronal activations, while calcium traces provide a full temporal
1102 picture of these activations together with baseline activity.

1103 All cells identified using CNMF-E were then scored as neurons or not by a human scorer.
1104 Scoring was also done within CIATAH software in a Matlab GUI. Scoring was done while
1105 visualizing and considering a calcium activity trace, average waveform, a montage of the
1106 candidate cell's Ca2+ events, and a maximum projection of all cells on which the candidate cell
1107 was highlighted. The relative fluorescence ($\Delta F/F_0$) local maxima of each identified cell were
1108 considered calcium event times.

1109

1110 **Cell cross registration across sessions and within session**

1111 Validation and registration were completed as documented in Wirtshafter and Disterhoft[2].
1112 Briefly, videos underwent five rounds of registration using Turboreg image rotation[9] with the
1113 CIATAH software[8,11]. Background noise, axons, and dendrites were removed using an image
1114 binarization threshold of 40% of the images' maximum value. Cells were matched across
1115 sessions using a distance threshold of a maximum of five pixels, with a minimum 2-D correlation
1116 coefficient of 0.5. Sessions were aligned to session A(n), the last session in environment A.
1117

1118 **Place cell identification and computing spatial mutual information**

1119 Place cells were identified using mutual information computed when the animals were
1120 running at speeds greater than or equal to 4cm/s. MI was computed for all cells; there was no
1121 calcium event rate criterion for included cells. To be considered significant, the computed mutual
1122 information (MI) must be greater than 95% of MI scores computed 500 times from shuffled
1123 positions[12]. To compute the MI for each cell, the training environments were divided into 2.5cm
1124 x 2.5cm bins. The calcium event rate of each cell and the occupancy of the animal were found
1125 for each bin. Rate and occupancy were smoothed with a Gaussian kernel with filter width of 3cm
1126 and Sigma of 0.5cm. Mutual information was computed during periods of movement as
1127 follows[2,12,13]:

1128 $$p = \frac{P_s}{P_o}$$
1129 $$M_s = \sum P_s$$
1130 $$M_o = \sum P_o$$
1131 $$MI = \sum p * \log_2\left(\frac{p}{M_s * M_o}\right)$$
1132 where:
1133 $$P_s = \text{calcium event probability in each bin}$$
1134 $$P_o = \text{occupancy probability at each bin}$$
1135

Mutual information using calcium traces was computed as above, except instead of P_s being calcium event probability per bin, the value of P_s was the average value of calcium trace in the bin.

We computed MI using both calcium events and calcium trace data. There was no significant difference between the number of place cells detected using calcium event data and calcium trace data, (paired t-test $t(24)=1.01$, $p>0.05$). Note that all place cell and place field measurements are presented with conditioning periods included, as the animal was frequently moving during conditioning periods. We also computed results while excluding conditioning periods and found no significant differences.

Computing CSUS mutual information

The computation of CSUS mutual information was very similarly to that for spatial mutual information. A 1.3 second period beginning at the start of the CS tone was either divided into 2 bins (CSUS-MI2) or 5 bins (CSUS-MI5) (Fig. 3c-3d). Mutual information was then computed using the following:

$$p = \frac{P_s}{P_o}$$

$$M_s = \sum P_s$$

$$M_o = \sum P_o$$

$$MI = \sum p * \log_2\left(\frac{p}{M_s * M_o}\right)$$

Where:

$$P_s = \text{calcium event probability in each CSUS bin}$$

$$P_o = \text{probability of individual CSUS occuring out of all CSUS bins}$$

Mutual information using calcium traces was computed as above, except that P_s did not represent the calcium event probability per bin, but the average value of calcium trace within the bin.

Remapping quantification

The place cell center was defined as the occupancy-normalized location with maximum number of calcium events while the animal was moving at 5cm/s or faster. Position was binned into 2.5cm square bins. The place cell centers at environments A and B, as well as at environment A across days and at environment B across days, were used to align each environment across days, as well as to align environment A to environment B.

Population vector correlation was calculated between two environments using calcium event data. Neurons present in both datasets (such as sessions A(n) and A(n-1), or A(n) and B(1)) were identified and their calcium event times were converted to rates using 0.75 second binning. These firing rates were then normalized using z-score normalization across each neuron's activity across time. The mean calcium event rate for each neuron in each environment was then computed. The correlation among population vectors of firing rates in the two environments was cpmputed:

$$population\ vector\ correlation = \frac{\sum_{i=1}^{N}\left(r_{\{i,A\}} - \overline{\{r\}}_A\right)\left(r_{\{i,B\}} - \overline{\{r\}}_B\right)}{\sqrt{\sum_{i=1}^{N}\left(r_{\{i,A\}} - \overline{\{r\}}_A\right)^2}\sqrt{\sum_{i=1}^{N}\left(r_{\{i,B\}} - \overline{\{r\}}_B\right)^2}}$$

Where:

$r_{\{i,A\}}$ and $r_{\{i,B\}}$ = firing rates of neuron *i* in environments A and B

$\overline{\{r\}}_A$ and $\overline{\{r\}}_B$ = mean firing rates across neurons in environments A and B

The factors in the denominator correspond to the standard deviation of the components of each population vector relative to their mean, computed in each environment.

Use of CEBRA versus alternative methods

We explored multiple different methods before settling on the use of CEBRA for this study. A short summary of each tested method can be found below:

- **Principal Component Analysis (PCA)**[14]: Principal component analysis (PCA) is a statistical method used to reduce the dimensionality of data while retaining as much variability as possible. This linear technique identifies the axes (principal components) in the dataset that maximize variance. The first principal component explains the most variance, the second explains the second most, and so on. Principal components are combinations of original features and may not always have clear or intuitive meanings. In agreement with previous hippocampal data[15], PCA required upwards of 15-25 components to capture 95% of the variance of the data. In addition, across and within all sessions and representations (spatial and task representations), the manifolds spanned by the largest PCs remained highly similar, with small principal angles in pairwise comparisons. This similarity in the orientation of the leading subspaces suggested that PCA did not distinguish between spatial or behavioral components of the task (Figure S5).
- **Independent Component Analysis (ICA)**[16]: Independent Component Analysis (ICA) is a computational technique used to separate a multivariate signal into additive, independent components. ICA operates under the assumption that observed data are linear mixtures of underlying, independent sources. It aims to find a linear transformation that maximizes the statistical independence of the estimated components. We found that ICA embeddings were unstable throughout the length of the recordings, and also did not clearly map onto behavioral states (Figure S6).
- **Isomap**[17]: Isomap is a manifold learning technique that seeks to capture the intrinsic geometric structure of data. Isomap is useful when linear methods like PCA cannot capture the intrinsic structure of the data, as it preserves the geodesic (curved) distances

in the reduced dimensionality space. Unlike linear methods such as PCA, Isomap can capture nonlinear relationships in the data. Interestingly, using Isomap, only about 5 neural modes were required to achieve a residual variance of 5-10%. However, the embedding shape did not relate to any discernable property of neural data or behavior (Figure S7). Dimensionality reduction was achieved, but the resulting representations were not interpretable (Figure S7).

- **MIND**[15,18]: MIND is a decoding method designed for integrating multiple data modalities to predict various features, particularly sensory and motor functions. MIND uses recurrent neural networks whose hidden variables provide a memory mechanism for remembering previous inputs; this approach is particularly apt for the analysis of time series data such as neural recordings. While MIND was very robust at distinguishing the different environments, it was not equipped to handle relatively short signals separated in time, such as the conditioning trials separated by intertrial intervals. Our analyses using MIND resulted in poor and unstable embeddings that could not be analyzed (Figure S8).

CEBRA[19] was chosen for this project for its ability to capture nonlinear relationships in the data and to create stable embeddings over short and long time periods. Additionally, spatial separations of components were well isolated and correlated well with observed behaviors.

Use of CEBRA for position decoding

Optimal parameters for decoding the position of each animal from neural activity were determined using an extensive grid search across learning rate, temperature, and number of iterations (Figure S9). Models created to compare different sessions, such as a model trained on data from session A(n) used to decode session B(1), were only trained on cells that occurred in both sessions. Models were trained on spike traces of these cells, labeled with the animal's (X,Y)

position. In all cases, 75% of data was used to train the model while 25% of data was held out for verification. All models were run 500 times. Optimal embeddings were determined based on the minimum median absolute error between the predicted and true positions. The optimal parameters for each rat are as follows:

	Rat 1	Rat 2	Rat 3	Rat 4	Rat 5
Model Architecture	'Offset10-model'				
Batch size	512				
Learning rate	$5.5*10^{-5}$	$6.625*10^{-4}$	$5.5*10^{-4}$	$1.0*10^{-3}$	$1.0*10^{-3}$
Temperature mode	'Auto'				
Minimum temperature	No minimum	1.5	0.95	$1.0*10^{-9}$	No minimum
Output dimensions (# of latents)	3				
Max iterations	25000	8000	26500	30000	18000
Distance	'Cosine'				
Conditional	'Time delta'				
Number of hidden units	32				
Time offsets	1				

The number of output dimensions was chosen based on the fewest number of dimensions under which all 5 models consistently outperformed shuffled data for both position and conditioning decoding (Figure S10-12).

Use of CEBRA for conditioning decoding

As in position decoding, the optimal parameters for decoding conditioning were determined for each animal using an extensive grid search across learning rate, temperature, and number of iterations (Figure S11). Models created to compare different sessions of neural activity, such as a model trained on data from session A(n) used to decode session B(1), were

only trained on cells that occurred in both sessions. Models were trained on spike traces of these cells, with labels corresponding to the CSUS bin during which the signal occurred (either one out of 2 bins or out of 5 bins, see Figures 3c-d). In all cases, 75% of data was used to train the model while 25% of data was held out for verification. All models were run 500 times. Optimal embeddings were determined based on the percent of correctly binned time points. The optimal parameters for each rat are as follows:

	Rat 1	Rat 2	Rat 3	Rat 4	Rat 5
Model Architecture	'Offset10-model'				
Batch size	512				
Learning rate	$3.5*10^{-3}$	$7.0*10^{-3}$	$3.5*10^{-3}$	$7.5*10^{-3}$	$9.5*10^{-3}$
Temperature mode	'Constant'	'Constant'	'Auto'	'Constant'	'Constant'
Minimum temperature	2.33	1.75	1.67	1.67	2.66
Output dimensions (# of latents)	3				
Max iterations	50000	7500	20000	18000	25000
Distance	'Euclidian'	'Cosine'	'Cosine'	'Euclidian'	'Cosine'
Conditional	'Time delta'				
Number of hidden units	32				
Time offsets	1				

The number of output dimension was chosen based on the fewest number of dimensions under which all 5 models consistently outperformed shuffled data for both position and conditioning decoding (Figure S10-12). The parameters listed above were used for decoding into 2 or 5 bins, including the use of 3 output dimensions (# of latents).

The accuracy of results was computed from the entries in the confusion matrix:

$$Accuracy = \frac{Sum\ of\ number\ of\ diagonal\ (correct)\ elements}{Sum\ of\ number\ of\ total\ elements}$$

Precision was calculated for each class i, $1 \leq i \leq n$, where n is the number of classes:

$$Precision_i = \frac{(TP_i)}{TP_i + FP_i}$$

where:

$$TP = true\ positives$$
$$FP = false\ positives$$

The global precision is given by the average:

$$Precision = \frac{1}{n}\sum_{i=1}^{n} Precision_i$$

Recall, also known as sensitivity, was calculated for each class i, $1 \leq i \leq n$, where n is the number of classes:

$$Recall_i = \frac{(TP_i)}{TP_i + FN_i}$$

where:

$$TP = true\ positives$$
$$FN = false\ negatives$$

The global recall is given by the average:

$$Recall = \frac{1}{n}\sum_{i=1}^{n} Recall_i$$

The F1 score was calculated for each class i, $1 \leq i \leq n$, where n is the number of classes:

$$F1_i = 2 * \frac{Precision_i * Recall_i}{Precision_i + Recall_i}$$

The global F1 score is given by the average:

$$F1 = \frac{1}{n}\sum_{i=1}^{n} F1_i$$

The area under the receiver operating characteristic (ROC) curve was calculated for each class i, $1 \leq i \leq n$, where n is the number of classes:

$$AUC_i = roc_auc_score(y_{true-bin}[:, i], y_{pred-prob}[:, i])$$

The global value is given by the average:

$$ROC\ AUC = \frac{1}{n}\sum_{i=1}^{n} AUC_i$$

where:

$$y_{true-bin} = binarized\ true\ labels$$
$$y_{pred-prob} = predicted\ probabilities\ for\ each\ class$$

Model consistency

Model consistency was computed using a built-in CEBRA function which relies on the function 'sklearn.metrics.consistency_score'[20]. The function compares the embeddings from different models by calculating pairwise consistency scores. This comparison involves measuring the similarity of the embeddings using statistical metrics; i.e. this metric calculates how similar a model's labels are for similar instances in the data set.

To determine consistency between environments and across animals, the data was fit to each model 20 times. The model with the lowest loss was selected and compared to other models with the lowest loss. Models were created using each individual animal's optimal parameters (see above).

Acknowledgements: This work was supported by an NIA T32 (T32-AG020506/AG/NIA), an NIA R37 (R37-AG008796/AG/NIA), an NINDS R01 (R01 NS113804/NS/NINDS), and a K99 award (K99 MH135062).

This research was supported in part through the computational resources and staff assistance provided by the Quest high performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

We would like to thank all members of the Disterhoft lab, especially Mackenzie Kneisly. We would like to thank the following individuals for their assistance with CEBRA: Mackenzie Mathis and Steffen Schneider. And additional thank you to David Wirtshafter for his feedback.

Author contributions: Investigation, H.S.W.; Formal Analysis, H.S.W.; Writing –Original Draft, H.S.W.; Writing – Review & Editing, H.S.W., S.A.S., J.F.D.; Funding: H.S.W, J.F.D., Supervision, S.A.S., J.F.D.

Declaration of interests: The authors declare no competing interests.

Additional Information

Supplementary Information is available for this paper.

Correspondence and requests for materials should be addressed to Hannah S Wirtshafter, hsw@northwestern.edu .

Reprints and permissions information is available at www.nature.com/reprints.

Supplementary figures

Figure S1

Figure S1. Photos of testing chambers.

Photos of testing chambers. Top: Environment A, an unscented rectangular enclosure with wire floor and walls, and white lighting. Bottom: Environment B, a scented ovular enclosure with white solid floor and walls, and red lighting. Both environments were located at the same spot in the room relative to external cues. The door to the chamber was closed during testing, to accentuate the distinction between the white and red lighting.

Figure S2

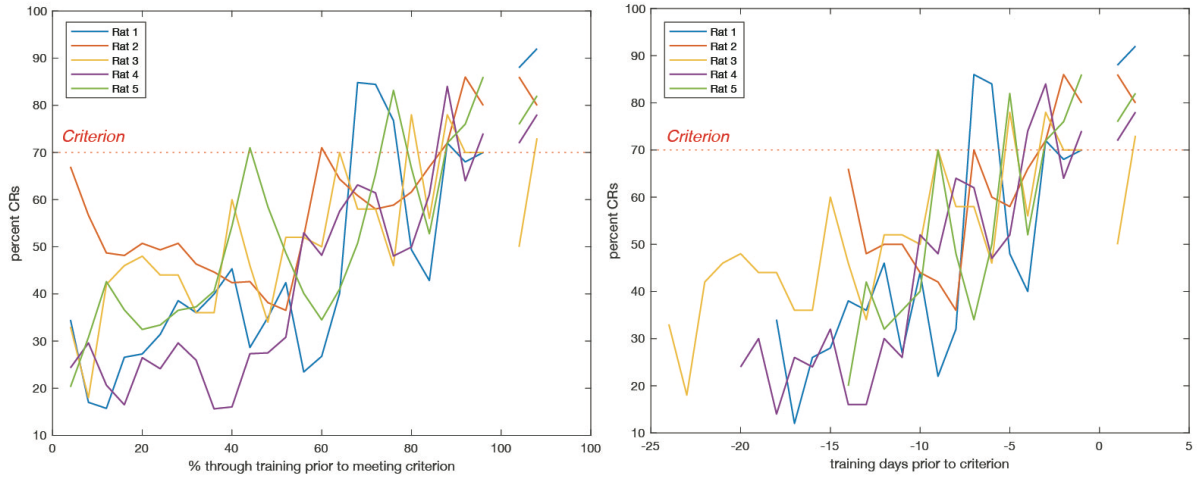


Figure S2. Learning curves for the five rats.

Substantial variability in the number of sessions required to learn the task; average number of sessions was 20 ± 4.2 (including the criterion sessions). The fastest learners (2 rats) reached criterion after 14 sessions, while the slowest rat required 24 sessions to reach criterion.

Figure S3

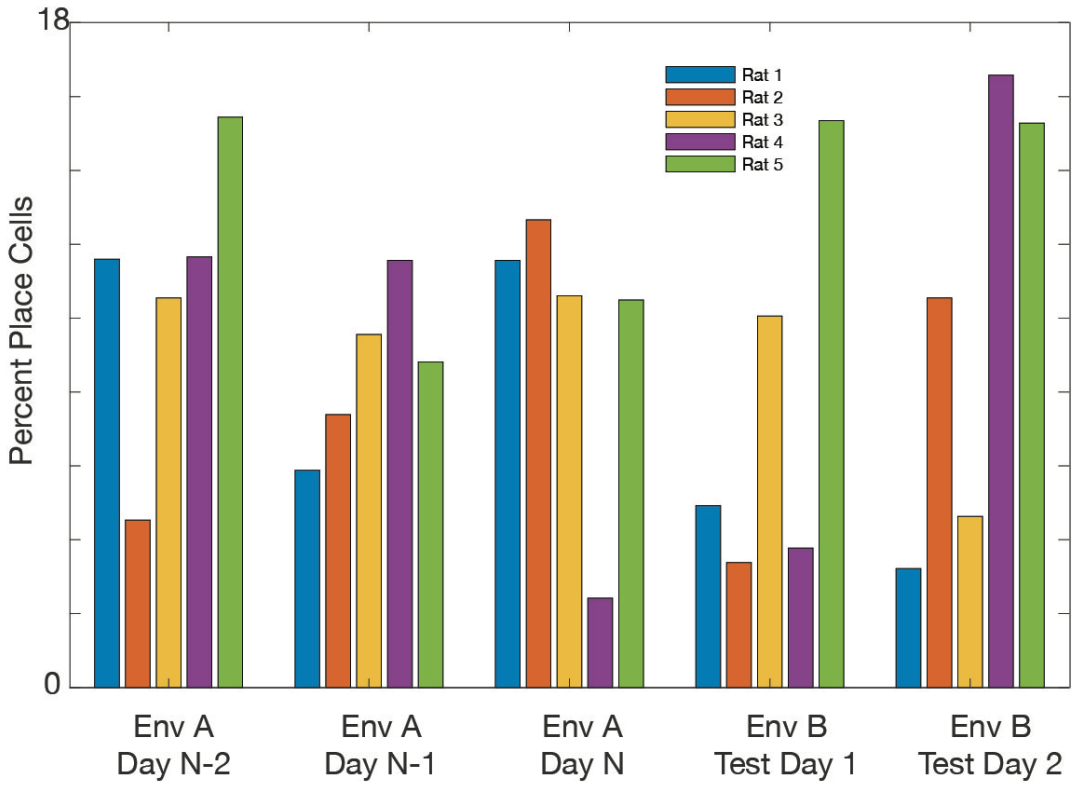


Figure S3. Percent of place cells by session for each animal.

The percentage of cells classified as place cells for each session and each animal

Figure S4

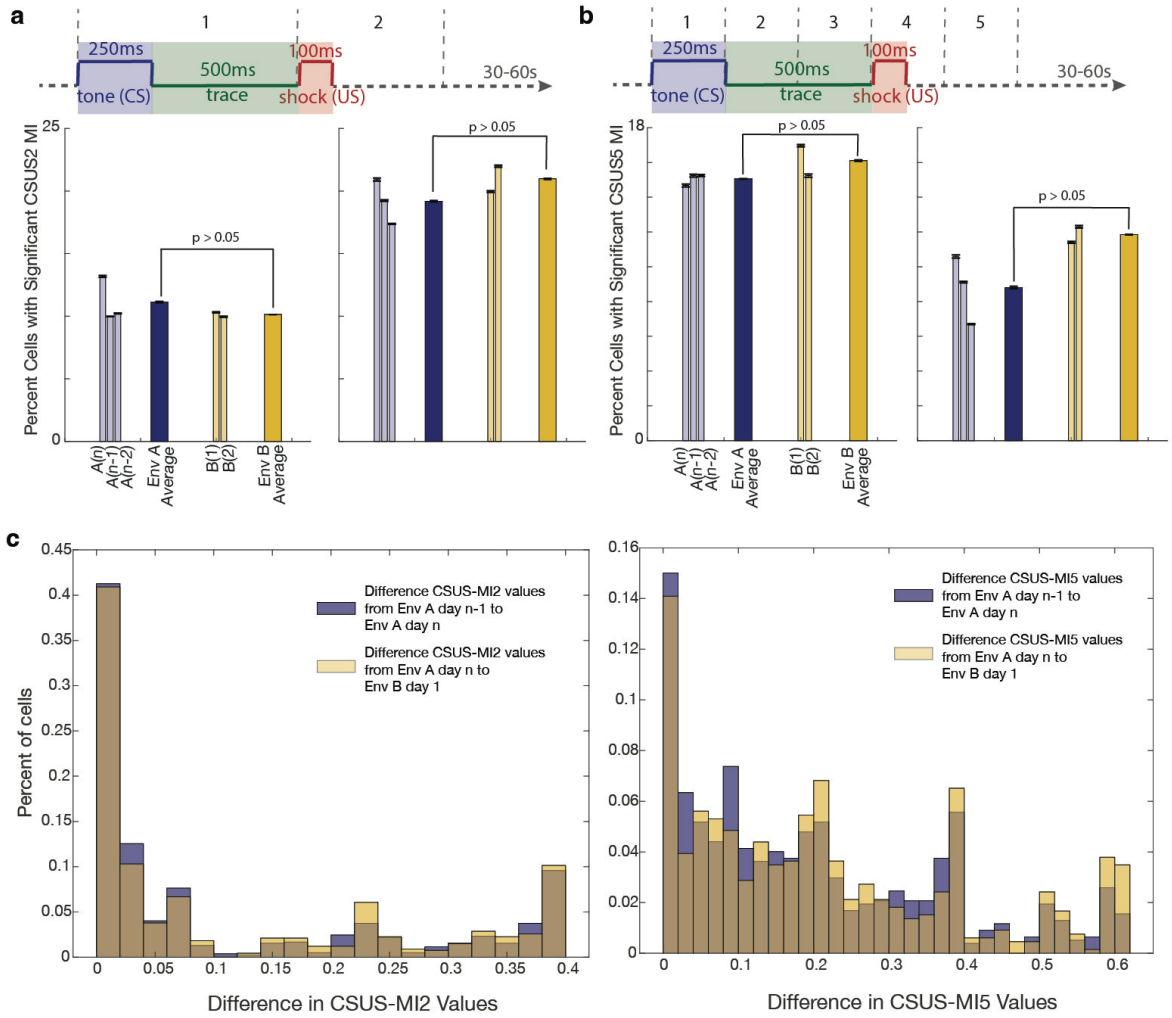


Figure S4. CSUS-MI2 and CSUS-MI5 differences between sessions.

a. The trial period was divided into two segments: the CS and trace period (750 ms) and the US and post-US period (500 ms). Mutual information (MI) was calculated for cells based on these two periods and compared to shuffled data, where period IDs were shuffled 500 times across all trials. Left: Using calcium event data, we found that 10.7% ± 4.9% of cells contained significant CSUS information related to whether the animal was in a CS or US period. Right: Using calcium traces, 19.9% ± 8.2% of cells contained significant information distinguishing the CS from the US period. No significant differences in

CSUS-MI were observed between environments A and B (double-sided t-tests, calcium events: $t(23) = 0.48$, $p > 0.05$; calcium traces: $t(23) = -0.52$, $p > 0.05$).

b. The trial period was divided into five equal segments, each 250 ms. MI was calculated for each cell based on these five periods. We compared the observed values to those obtained after shuffling period IDs 500 times. Left: Using calcium event data, 15.5% ± 7.8% of cells contained significant information distinguishing the five periods, compared to 10.0% ± 7.8% when using calcium trace data. No significant differences in these MI metrics were found between environments A and B (double-sided t-tests, calcium events: $t(23) = -0.32$, $p > 0.05$; calcium traces: $t(23) = -1.1$, $p > 0.05$).

c. Left: There was no significant difference in CSUS-MI2 values when comparing session A(n) to session A(n-1) versus session A(n) to session B(1) (Wilcoxon rank sum test: $p > 0.05$; double-sided t-test: $t(1431) = 0.86$, $p > 0.05$). Right: A small but significant difference was observed in CSUS-MI5 when comparing session A(n) to session A(n-1) versus session A(n) to session B(1) (Wilcoxon rank sum test: $p = 0.049$; double-sided t-test: $t(1431) = -2.2$, $p = 0.03$).

Figure S5

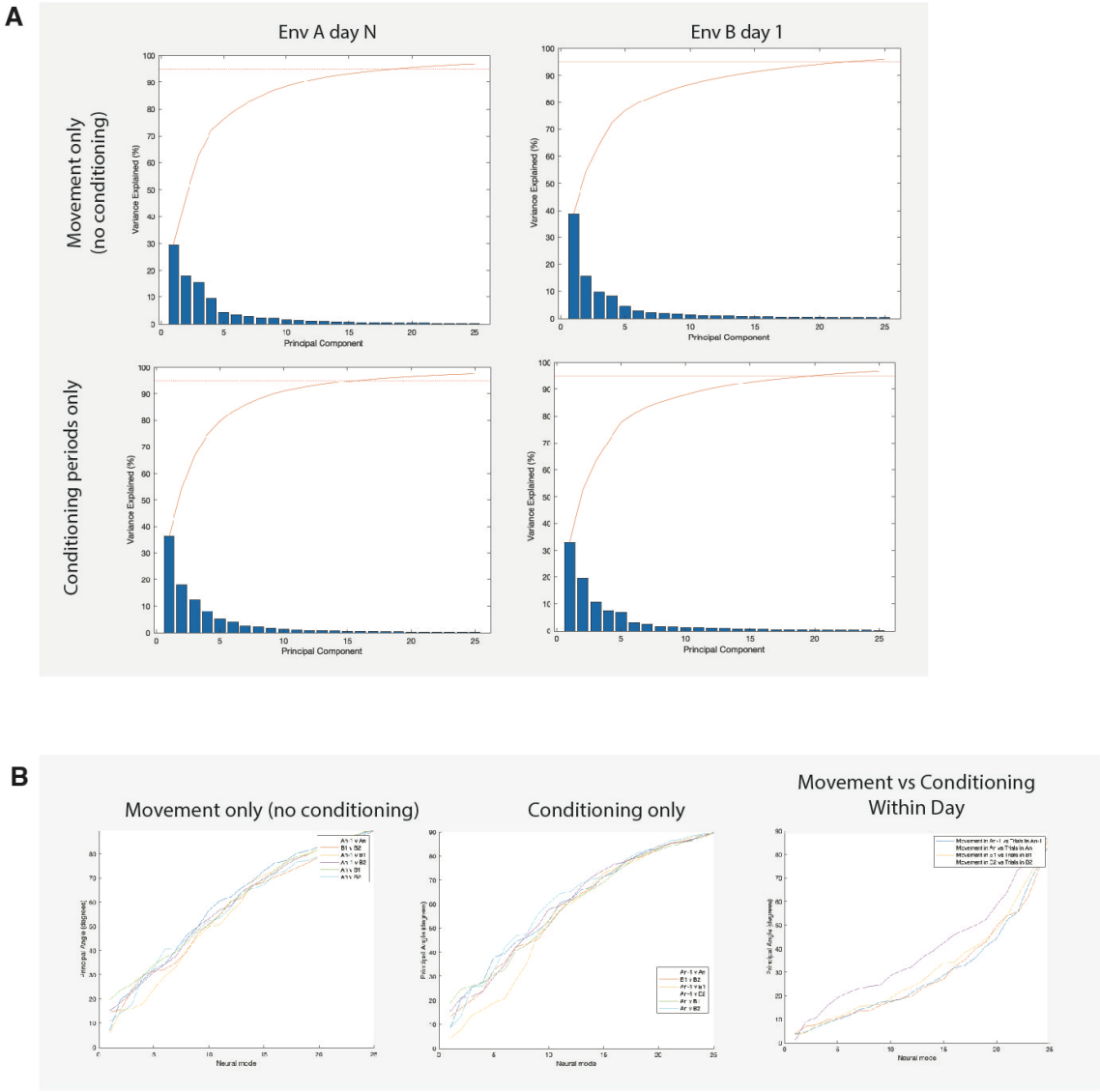


Figure S5. PCA computations for session A(n) and session B(1).

Only cells present in both sessions were used. Principal component analysis (PCA) revealed that approximately 15-25 principal components (PCs) are needed to account for 95% of the variance in the data. When using the complete cell population (not shown), more than 25 PCs are required to achieve the same variance. Across and within all sessions and representations (spatial and task), the principal angles between manifolds remain highly similar.

Figure S6

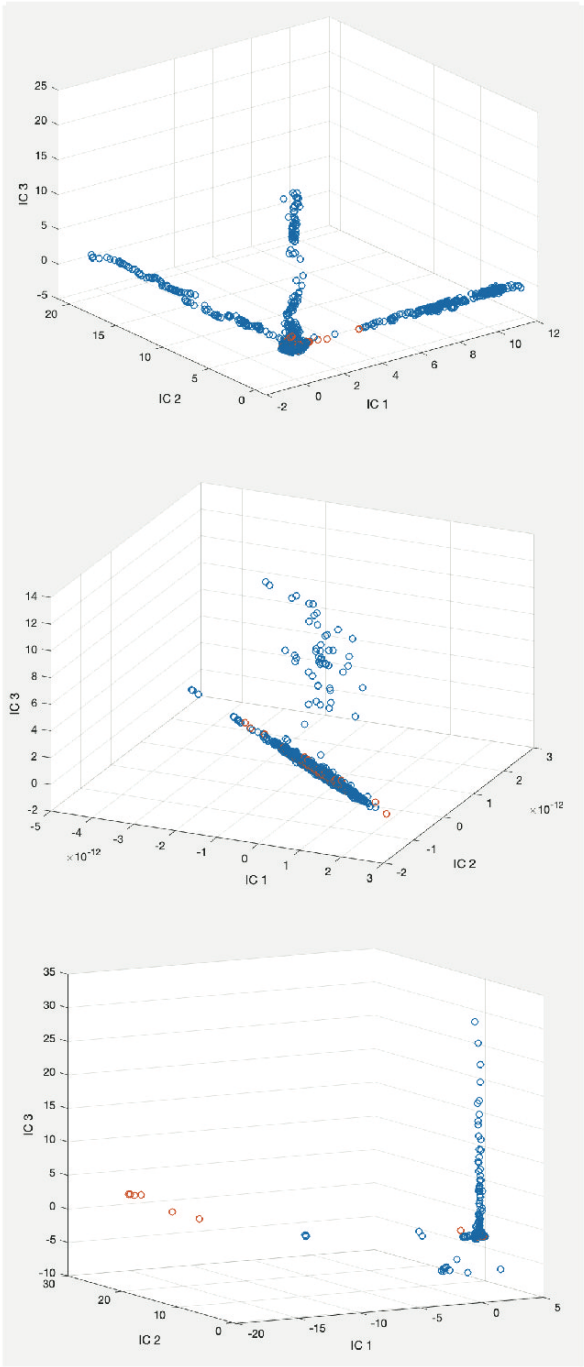


Figure S6. ICA computations across different segments of a session.

Top: Independent component analysis (ICA) was for data over an entire session, using three independent components (ICs). Blue dots represent non-trial times, while red dots represent trial times. Middle: ICA computed over the last two-thirds of the same session shows variability in

ICs across segments within a session. Bottom: ICA computed over the second half of the session shows additional variability in the results of the analysis depending on how the session is divided. These results indicate that independent components are highly variable over the course of one session and are sensitive to how the session is partitioned.

Figure S7

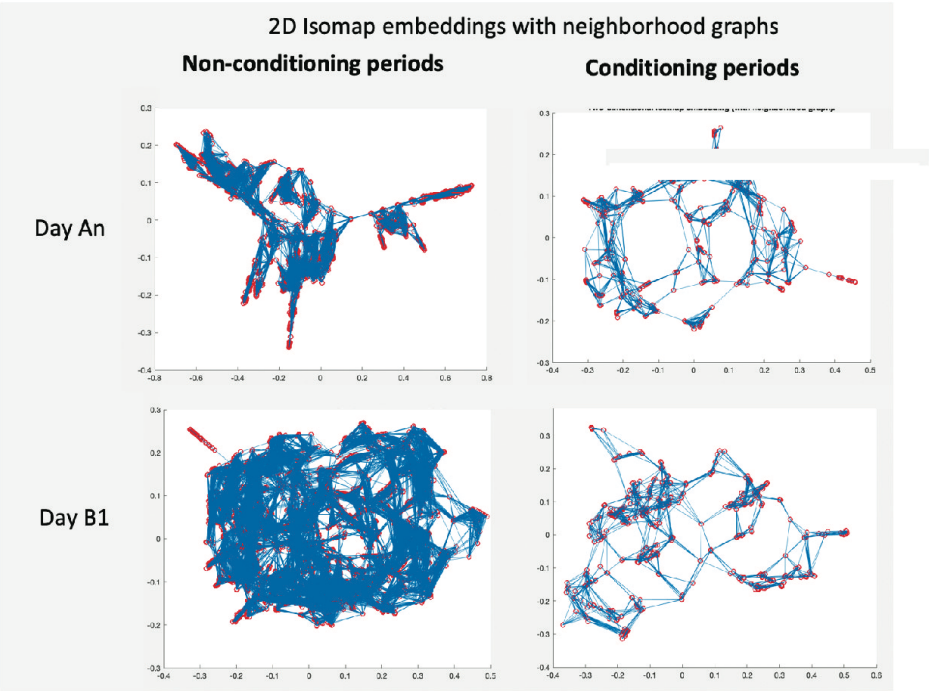


Figure S7. Isomap computations for session A(n) and session B(1).

Isomap computations suggest that approximately five neural modes are sufficient to achieve a residual variance of 5-10%. However, the shape of the Isomap embedding does not correlate with any discernable properties of neural activity or behavior, suggesting limited interpretability of the embedding structure.

Figure S8

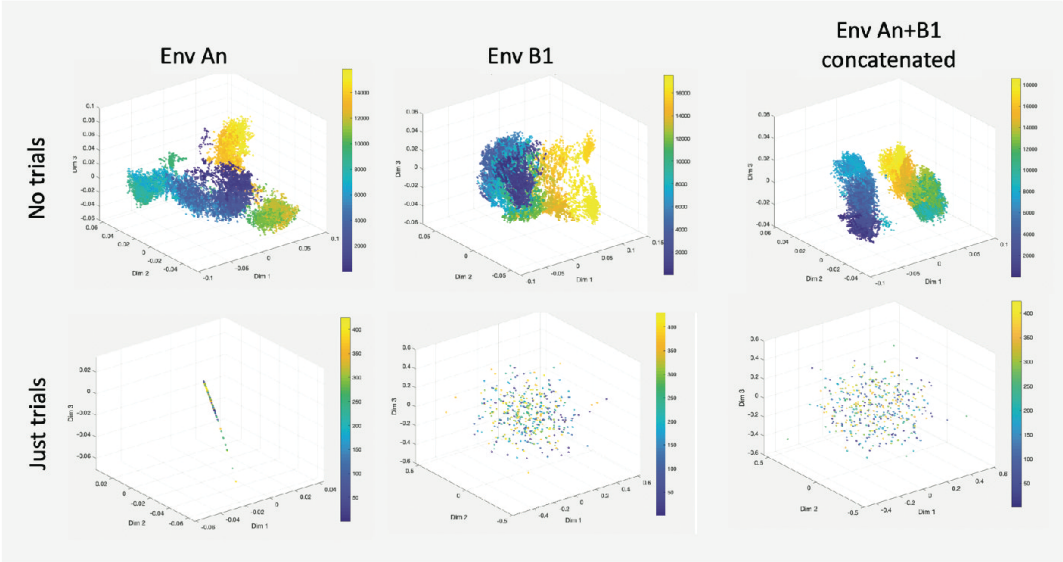


Figure S8. MIND outputs for sessions A(n), B(1), and concatenated sessions. Top row: MIND embeddings during movement, excluding trial periods, with color bars representing frames. The temporal structure of the data is well captured, with clear separation between A(n) and B(1). Bottom row: MIND embeddings during conditioning periods are highly unstable. Small changes in parameters result in substantial shifts in the embedding structure, transitioning from a linear structure (left) to an undefined, unstable cloud (middle and right).

Figure S9

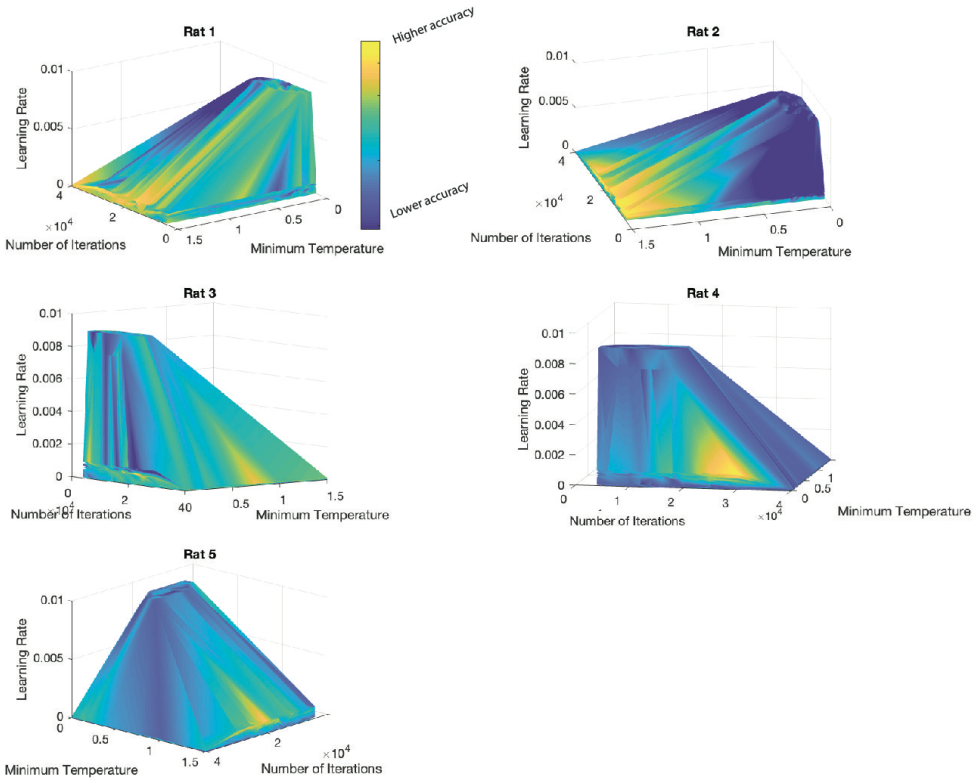


Figure S9. Grid search over decoding parameters for position.

A grid search was performed over three parameters: minimum temperature, learning rate, and number of iterations, for decoding position. Models were trained using cells from session A(n) that also appeared in session A(n-1). The figure shows decoding accuracy for session A(n-1) using the models trained on data from session A(n). Yellow areas indicate higher decoding accuracy.

Figure S10

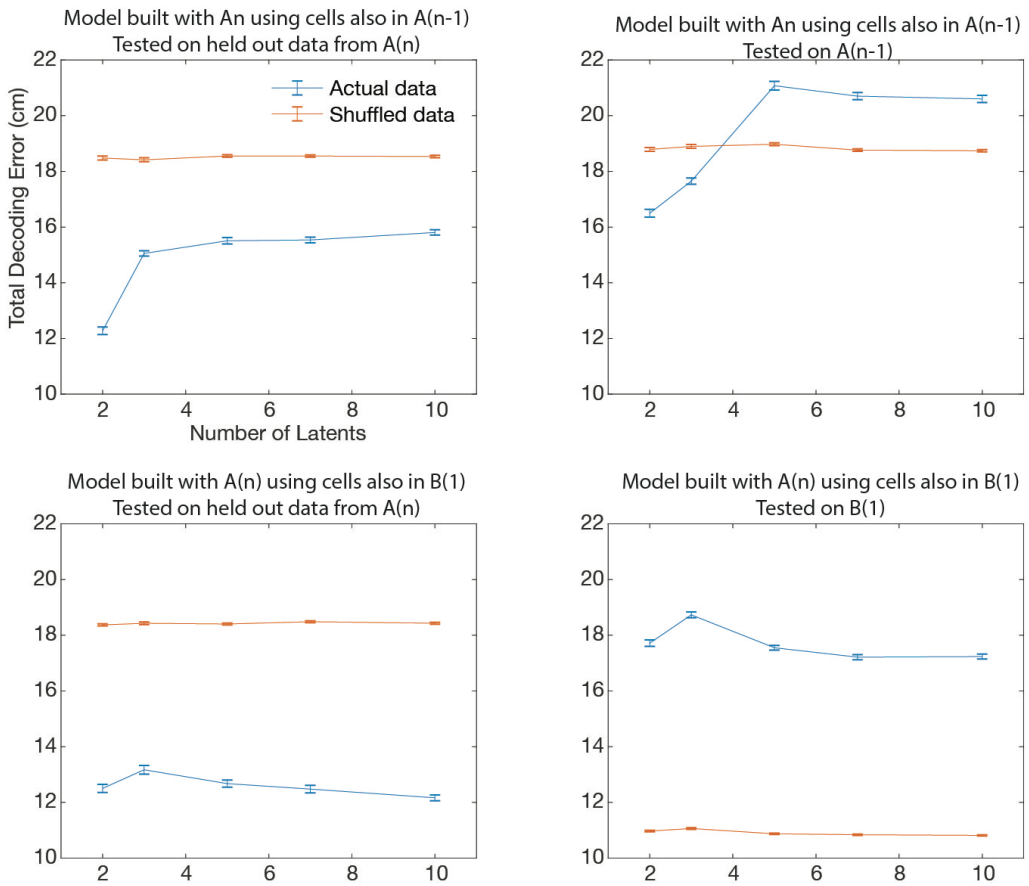


Figure S10. Position decoding error as a function of the number of latents (Rat 5). This figure shows the decoding error for position as the number of latents increases. Left panels: model trained on data from A(n), tested on held out data from A(n). Upper left panel: As the number of latents increases, the model's ability to decode a different session within the same environment decreases. This effect is not consistent across all rats; see lower left panel. Right panels: model trained on data from A(n), tested on A(n-1) (top) or B(1) (bottom). Performance is particularly bad when the model is tested on a different environment (lower right panel).

Figure S11

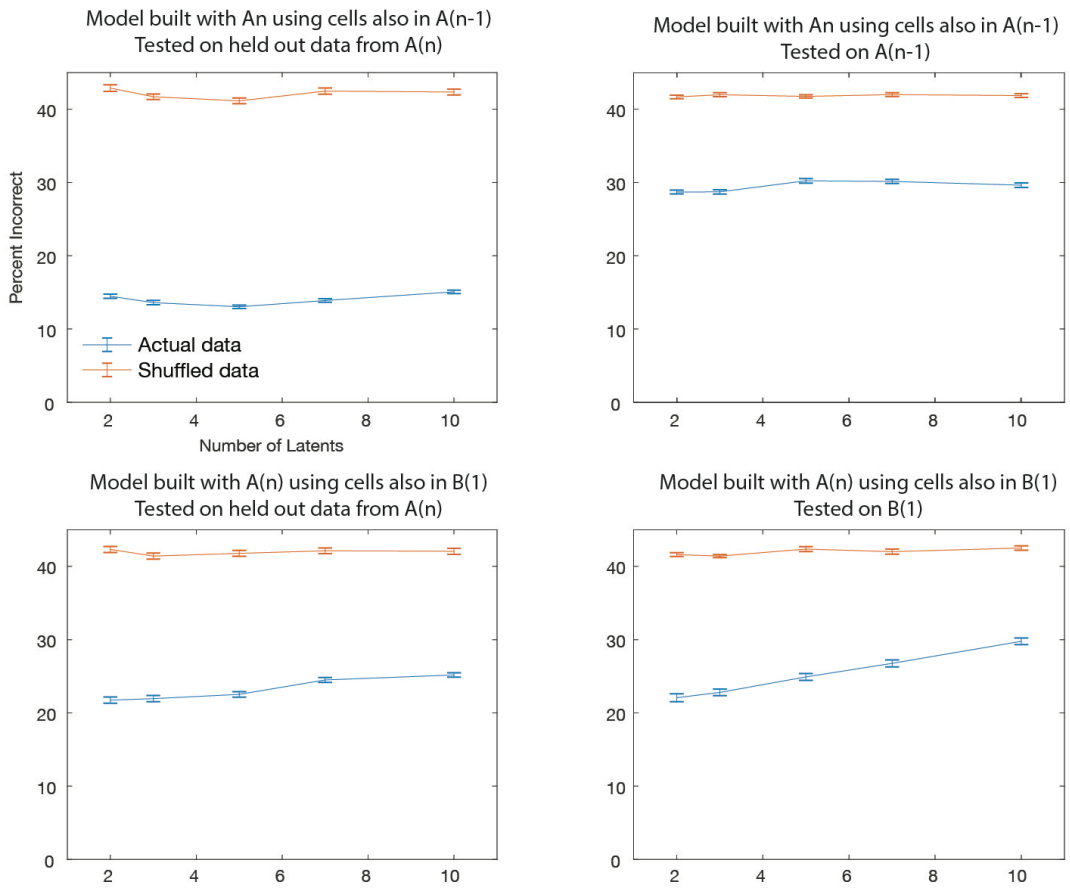


Figure S11. CSUS2 decoding accuracy with increasing number of latents (Rat 3). Percent of incorrect decoding for the CSUS2 model as the number of latents increases. Left panels: model trained on data from A(n), tested on held out data from A(n). As the number of latents increases, the model's ability to decode a different session within the same environment remains stable or slightly decreases. Right panels: model trained on data from A(n), tested on A(n-1) (top) or B(1) (bottom). Performance remains stable for a different session in the same environment (upper right panel) but deteriorates with increasing number of latents when the model is tested in a different environment (lower right panel). Each model was run 100 times.

Figure S12

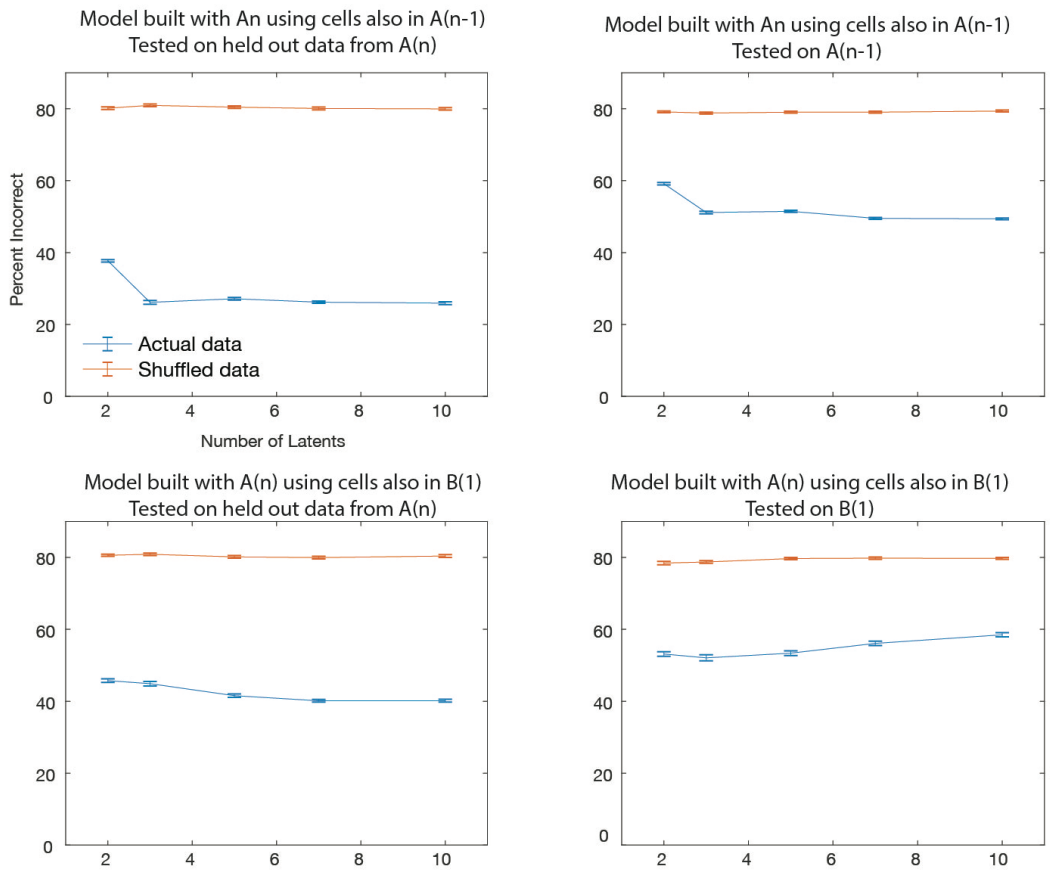


Figure S12. CSUS5 decoding accuracy with increasing number of latents (Rat 5). Same as Figure S11, but for CSUS5; the conditioning period has been divided into five time bins instead of two. The percent of incorrect decoding is shown as the number of latents increases. Left panels: model trained on data from A(n), tested on held out data from A(n). As the number of latents increases, the model's ability to decode a different session within the same environment remains stable or slightly increases. Right panels: model trained on data from A(n), tested on A(n-1) (top) or B(1) (bottom). Performance remains stable for a different session in the same environment (upper right panel) but deteriorates with increasing number of latents when the model is tested in a different environment (lower right panel). Each model was run 100 times.

Figure S13

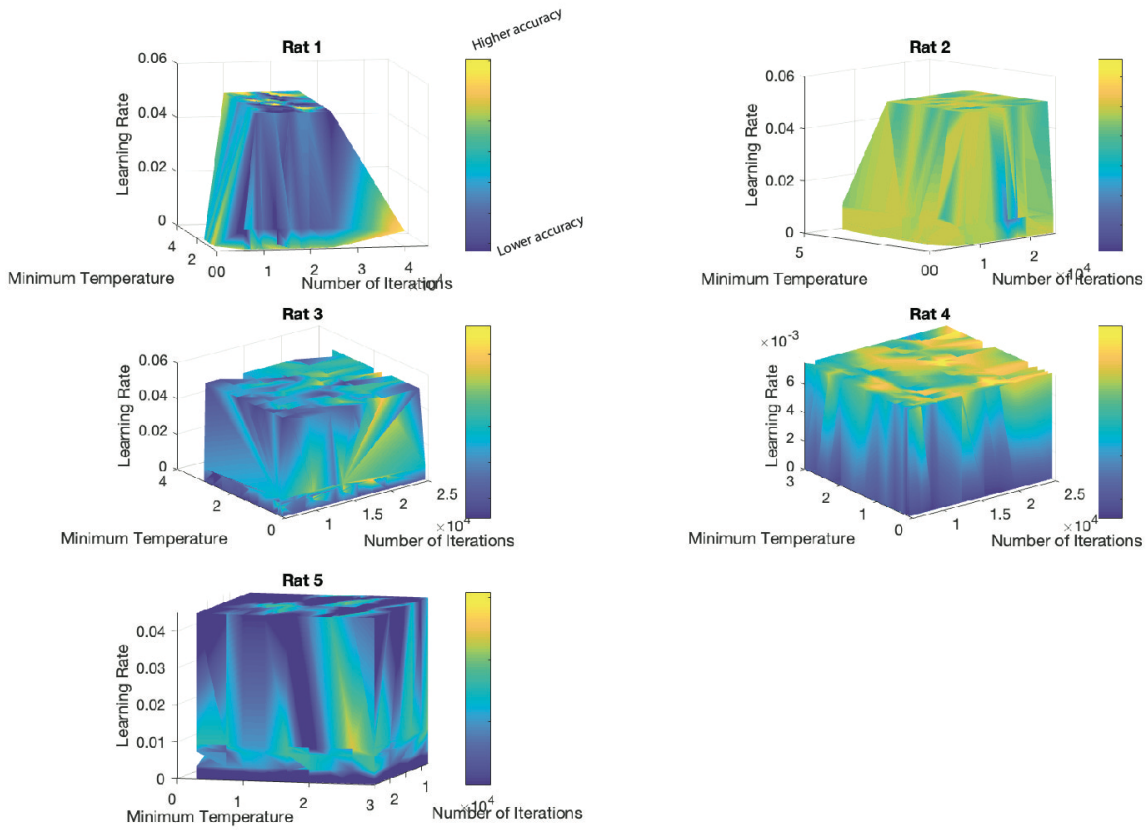


Figure S13. Grid search over decoding parameters for conditioning.

A grid search was performed over three parameters: minimum temperature, learning rate, and number of iterations for conditioning decoding. Models were trained using cells from session A(n) that also appeared in session B(1). The figure shows decoding accuracy for CSUS2 session B(1) using the models trained on data from session A(n). Yellow areas indicate higher decoding accuracy. For Rats 1 and 4, the 'Euclidean' distance with 'constant' temperature mode was used. For Rats 2 and 5, 'cosine' distance with 'constant' temperature mode was used. For Rat 3, 'cosine' distance with 'auto' temperature mode was used.

Methods References

1 Wirtshafter, H. S. & Disterhoft, J. F. Place cells are nonrandomly clustered by field location in CA1 hippocampus. *Hippocampus* **33**, 65-84 (2023). https://doi.org:10.1002/hipo.23489

2 Wirtshafter, H. S. & Disterhoft, J. F. In Vivo Multi-Day Calcium Imaging of CA1 Hippocampus in Freely Moving Rats Reveals a High Preponderance of Place Cells with Consistent Place Fields. *J Neurosci* (2022). https://doi.org:doi: 10.1523/JNEUROSCI.1750-21.2022

3 Weiss, C. & Thompson, R. F. The effects of age on eyeblink conditioning in the freely moving Fischer-344 rat. *Neurobiol Aging* **12**, 249-254 (1991). https://doi.org:10.1016/0197-4580(91)90105-s

4 Skelton, R. W. Bilateral cerebellar lesions disrupt conditioned eyelid responses in unrestrained rats. *Behav Neurosci* **102**, 586-590 (1988). https://doi.org:10.1037//0735-7044.102.4.586

5 Aharoni, D., Khakh, B. S., Silva, A. J. & Golshani, P. All the light that we can see: a new era in miniaturized microscopy. *Nat Methods* **16**, 11-13 (2019). https://doi.org:10.1038/s41592-018-0266-x

6 Silva, A. J. Miniaturized two-photon microscope: seeing clearer and deeper into the brain. *Light: Science & Applications* **6**, e17104-e17104 (2017).

7 Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience* **21**, 1281-1289 (2018).

8 Corder, G. *et al.* An amygdalar neural ensemble that encodes the unpleasantness of pain. *Science* **363**, 276-281 (2019). https://doi.org:10.1126/science.aap8586

9 Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process* **7**, 27-41 (1998). https://doi.org:10.1109/83.650848

10 Zhou, P. *et al.* Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *eLife* **7** (2018). https://doi.org:10.7554/eLife.28728

11 CIAtah: a software package for analyzing one- and two-photon calcium imaging datasets. v. v1.0.0 (Zenodo, 2018).

12 Kinsky, N. R., Sullivan, D. W., Mau, W., Hasselmo, M. E. & Eichenbaum, H. B. Hippocampal Place Fields Maintain a Coherent and Flexible Map across Long Timescales. *Current biology : CB* **28**, 3578-3588 e3576 (2018). https://doi.org:10.1016/j.cub.2018.09.037

13 Olypher, A. V., Lansky, P., Muller, R. U. & Fenton, A. A. Quantifying location-specific information in the discharge of rat hippocampal place cells. *J Neurosci Methods* **127**, 123-135 (2003). https://doi.org:10.1016/s0165-0270(03)00123-7

14 Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science* **2**, 559-572 (1901).

15 Nieh, E. H. *et al.* Geometry of abstract learned knowledge in the hippocampus. *Nature* **595**, 80-84 (2021).

16 Ans, B., Hérault, J. & Jutten, C. Architectures neuromimétiques adaptatives: Détection de primitives. *Proceedings of Cognitiva* **85**, 593-597 (1985).

17 Tenenbaum, J. B., Silva, V. d. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *science* **290**, 2319-2323 (2000).

18 Low, R. J., Lewallen, S., Aronov, D., Nevers, R. & Tank, D. W. Probing variability in a cognitive map using manifold inference from neural dynamics. *BioRxiv*, 418939 (2018).

19 Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360-368 (2023). https://doi.org:10.1038/s41586-023-06031-6

20 Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* **12**, 2825-2830 (2011).