



Published in final edited form as:

Nature. 2022 August ; 608(7921): 153–160. doi:10.1038/s41586-022-04936-2.

Prefrontal Feature Representations Drive Memory Recall

Nakul Yadav^{1,2}, Chelsea Noble^{1,†}, James E. Niemeyer^{1,†}, Andrea Terceros¹, Jonathan Victor², Conor Liston², Priyamvada Rajasethupathy^{1,*}

¹Laboratory of Neural Dynamics & Cognition, The Rockefeller University, New York, NY 10065 USA

²Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY 10065, USA

Summary

Memory formation involves binding of contextual features into a unitary representation^{1–4}, whereas memory recall can occur using partial combinations of these contextual features. As such, the neural basis underlying the relationship between a contextual memory and its constituent features is not well understood; in particular, where features are represented in the brain and how they drive recall. To gain insight into this question, we developed a behavioral task where mice use features to recall an associated contextual memory. We performed longitudinal imaging in hippocampus as mice performed this task and identified robust representations of global context but not of individual features. To identify putative brain regions that provide feature inputs to hippocampus, we inhibited cortical afferents while imaging hippocampus during behavior. We found that while inhibition of entorhinal cortex led to broad silencing of hippocampus, inhibition of prefrontal anterior cingulate led to a highly specific silencing of context neurons and deficits in feature-based recall. We next developed a preparation for simultaneous imaging of anterior cingulate and hippocampus during behavior, which revealed: 1. robust population-level representation of features in anterior cingulate, that 2. lag hippocampus context representations during training, but 3. dynamically reorganize to lead and target recruitment of context ensembles in hippocampus during recall. Together, we provide the first mechanistic insights into where contextual features are represented in the brain, how they emerge, and access long-range episodic representations, to drive memory recall.

A contextual memory is the unification of multiple streams of sensory information entwined in a spatiotemporal framework. The hippocampus encodes such memories as a unified conjunctive representation^{1–4}. This is thought to result from recurrent networks that

* Correspondence: priya@rockefeller.edu.

† These authors contributed equally

Author Contributions

NY and PR conceived the study and designed the experiments. NY and JN developed and characterized the VR-based behavioral task. NY performed imaging in hippocampus, anterior cingulate cortex (together with JN), paired inhibition and imaging experiments (together with CN), and dual imaging experiments (together with CN). NY performed data analyses with supervision from JV, CL, and PR. AT assisted with optogenetics experiments. NY and PR wrote the paper with input from all authors.

Declaration of Interests: The authors declare no competing financial interests.

Code Availability: Source extraction codes used in this study are publicly available. The custom analysis codes are available from the corresponding author upon request.

merge associated concepts into a singular representation across the CA3-CA1 network⁵⁻⁸. Thus, global conjunctive representations of a context are widely reported in hippocampus, but it remains unknown where individual features are stored. The existence of a feature representation, separate from a conjunctive representation, would allow features to be independently accessed during feature recognition^{9,10}, feature-based memory retrieval¹¹, memory updating^{12,13}, and adaptive coding^{14,15}. Yet, the neurobiological substrates of feature representations remain unclear.

It is possible that feature representations are 1) embedded within hippocampus, such that neurons encoding discrete sensory stimuli^{3,16,17} function as feature selective neurons capable of recruiting a population level contextual response, or 2) laid out in separate, even distributed, brain areas that have targeted access to hippocampal conjunctive representations. To address this question, we developed approaches to perform real-time visualization and manipulation of hippocampal and cortical circuits as mice form, store, and recall multi-modal contextual memories.

Conjunctive memory representations in CA1

We developed a head-fixed virtual reality-based memory retrieval task in which mice navigate an endless corridor and repeatedly experience three randomly sequenced multimodal contexts, each defined by a unique combination of sensory cues (A-auditory, V-visual, O-olfactory, and T-texture (Fig. 1a; and Methods). During training, mice were trained to associate one context with reward (sucrose delivery, leading to enhanced licking), another as neutral (water delivery), and a third context as aversive (airpuff and water delivery, leading to lick suppression). Learning was assessed by significant modulation of lick rates on 20% of interspersed probe trials where reinforcement was omitted.

During training, mice demonstrated a significant increase in lick rate (Fig. 1a, n=12 mice, 24 sessions) and decrease in latency to lick (Extended Data Fig 1a), on probe trials of reward vs. aversive context, indicating successful learning of the contextual associations. This learning was hippocampal dependent (Extended Data Fig 1c, Methods).

During the retrieval phase of the task, mice were provided with trials consisting of full (i.e., AVOT) or partial features (i.e., OT, AT, AOT) of the original multi-modal contexts in the absence of reinforcements (n=12 trials/day). On average, mice performed successful retrieval in full and partial feature trials (n=3 mice shown, Fig. 1b, Extended Data Fig. 1b). Interestingly, individual mice used different features to drive memory recall (Fig. 1b), suggesting the use of sensory integration rather than a single externally directed salient cue to learn and recall contextual memories. Of note, since some mice did not reliably classify neutral context as a distinct third context, in subsequent experiments we focused our analysis on reward and aversive contexts only.

To assess whether hippocampus encodes context representations, features, or both, mice were injected with a genetically encoded Ca²⁺ indicator, GCaMP6f, and implanted with a gradient index (GRIN) lens above dorsal CA1. This enabled volumetric two-photon imaging of the same field of view (FOV) throughout training and retrieval (Fig. 1c;

Extended Data Figs. 1d,e; 2a). Fast volumetric images ($800 \times 800 \mu\text{m}$ x/y, $150 \mu\text{m}$ z) were collected, providing access to >500 neurons per session (Extended Data Fig. 2b), from which significant calcium events were extracted (Extended Data Fig. 2c–g) and aligned to stimuli and behavior. During training, a substantial fraction (~10%) of neurons responded selectively to either the reward or aversive context (Fig. 1d,e,g). Importantly, these context selective neurons emerge only after learning (is not purely a sensory representation, Extended Data Fig. 3a,c), remain stably selective throughout retrieval (Extended Data Fig. 3b,d), exist on probe trials (is not due to reinforcement, Fig. 1d,e; Extended Data Fig. 3e), and are not time-locked to speed or licking (is not movement related, Extended Data Fig. 4a–e), and are thus a cognitive signal of context representing the learned association. At the population level, we also observed significant divergence of neural trajectories and high context decoding accuracy within the first few seconds of trial onset (Extended Data Fig. 3f,g).

During retrieval, we again observed robust conjunctive responses, but rarely observed neurons with feature-selective responses (Fig. 1f,g, 9% conjunctive vs. 1.1% feature-selective, chance= 0.5% and 0.8% respectively, Extended Data Fig. 3h; n=3 mice, 9 sessions). In fact, almost all neurons that responded to a particular feature of a given context (Methods, Extended Data Fig. 2g,h), also had significant and prominent responses to all other features of the same context (Fig. 1f; Extended Data Figs. 3b,d, 4f). At the population level, neural state-space trajectories projected onto a low dimensional subspace¹⁸, demonstrated prominent divergence of context-evoked trajectories but no discernable separation of feature evoked trajectories (Fig. 1h, Extended Data Fig. 4g). Finally, we assessed the performance of a linear support vector machine to classify context when trained on trials containing three different features and tested on trials of a fourth held-out feature. We found that the decoding accuracy of the model exceeded 80% within the first few second of trial onset for any feature type (Fig. 1i) indicating significant shared underlying dynamics for all features of a given context. Taken together, these data demonstrate a prominent conjunctive representation of multi-modal context in hippocampus, with lack of feature-specific representations, as previously suggested^{19–21}. This raises two questions 1) where are features encoded and 2) how do they interact with or recruit appropriate conjunctive representations in hippocampus for memory recall (Fig. 1j)?

AC, not LEC, required for feature-based recall

We suspected that extra-hippocampal brain circuits provide feature inputs to CA1 since intra-hippocampal circuits route through CA3 where the recurrent architecture is likely to create conjunctive rather than feature representations. We performed retrograde tracing to identify brain regions that send direct projections to dorsal hippocampus (Fig 2a; Extended Data Fig. 5a, 7f). Two of the strongest projections originated from the anterior cingulate region of the prefrontal cortex and lateral region of the entorhinal cortex, both of which we confirmed also synapse onto CA1 neurons in the anterograde direction (Extended Data Fig. 7a–c). Both regions have known roles in memory processing: We²² and others^{23–25} have previously demonstrated that prefrontal cortex is required for memory retrieval, and the identification of direct mono-synaptic prefrontal-to-hippocampus projections^{22,26} (Supplementary Note 1), highlight anterior cingulate (AC) cortex as an attractive candidate

for feature-recognition and recall. On the other hand, the hippocampus also receives dense projections from the lateral entorhinal cortex (LEC), where both conjunctive and feature-like responses have been observed^{27–30}. Thus, we aimed to dissect whether feature codes necessary for memory recall could be conveyed to the hippocampus from the AC, LEC, or both.

To dissect the contributions of each region to hippocampal physiology and behavior, we sought to inhibit AC or LEC while simultaneously imaging CA1 during feature based recall. We expressed a soma-localized inhibitory opsin, st-GtACR2³¹ in excitatory neurons of AC or LEC and implanted low-profile, angled fiber optic cannulas, ipsilaterally to the GRIN-imaged CA1 region. We targeted a field of view in CA1 known to contain inputs from both LEC and AC, (Fig. 2b; Extended Data Fig. 5b). During retrieval, light was delivered on half of the trials, while the other half served as light-off controls. *Light-off* trials achieved robust inhibition of post-synaptic neurons in CA1 (Fig. 2c,d). We aimed to minimize spectral cross-talk, by titrating 470nm light power at the fiber tip to <1.5mW to minimize unintended activation of GCaMP, and also confirmed that during two-photon imaging at 920nm, there was minimal unintended activation of soma-localized GtACR (Extended Data Fig. 6a–d, Methods). We found that Inhibition of AC led to silencing of a substantially greater proportion of context-selective neurons than did LEC inhibition, and than would be expected by chance, despite similar overall CA1 silencing elicited by both projections (Fig. 2d, Extended Data Fig. 6i).

To further characterize the functional inhibition in CA1, we quantified changes in dF/F onsets, or the fraction of time CA1 context neurons were active, in *light on vs light off* trials, compared to non-context neurons. We observed that LEC inhibition led to broad and widespread inhibition of CA1 neurons, whereas AC inhibition led to selective inhibition of context- neurons with negligible inhibition of non-context neurons (Fig. 2e,f). Quantifying this inhibition across all mice reveals a specificity for silencing of context- neurons during AC inhibition, for both full and partial-feature retrieval conditions, and in both reward and aversive contexts, which is absent in LEC inhibition (Fig. 2g; Extended Data Fig. 6e–h,j,k). Consistent with this, anterograde tracing from AC/LEC overlaid with cFos stain following training (to mark context neurons) revealed a greater fraction of CA1 neurons receiving inputs from AC compared to LEC (Extended Data Fig. 7d,e, n=4 mice each, p<0.01, Welch's *t*-test). In further support of this functional divergence between AC and LEC, we observed that AC and LEC project to almost entirely separate subset of neurons in CA1 (Extended Data Fig. 7a–c).

To determine whether the targeted suppression of CA1 context neurons by AC inhibition leads to behavioral deficits, we inhibited AC or LEC (CaMKII-stGtACR-mCherry) bilaterally (with CaMKII-mCherry as controls) and delivered light on half of the trials, while the other half served as *light-off* controls. We observed significant deficits in feature-based recall in *light-on* compared to *light-off* trials for all feature-based retrieval conditions during AC inhibition, which was absent during LEC inhibition, and in AC or LEC mCherry control mice. These data indicate a prominent role for AC in directing memory recall (Fig. 2h–i), and further that this occurs via targeted interactions with CA1 (Extended Data Fig. 7e, with AC sending direct excitatory inputs to CA1, Extended Data Fig. 7g–i) though we do not

exclude the possibility that these effects are also due to other downstream collaterals of AC or indirect interactions with CA1.

Feature representations in AC

To test whether AC indeed encodes for features, we next performed longitudinal two-photon imaging of AC throughout training and retrieval, targeting the sub- region of AC with known direct projections to hippocampus²². After injecting GCaMP6f in AC, we performed a craniotomy, and implanted a coverslip to gain chronic optical access to layers 2/3 of AC neurons (Fig. 3a; Extended Data Fig. 8a–h). Fast volumetric images ($800 \times 800 \mu\text{m x/y}$, $150 \mu\text{m z}$) of AC were collected providing access to >1000 neurons per session. During training, and similarly to CA1, AC neurons displayed robust context-selectivity at the single neuron and population levels (Fig. 3b, Extended Data Fig. 8i,j). However, during retrieval, we observed a stark difference in the tuning properties of single neurons in AC compared to CA1. AC neurons displayed feature-selectivity as well as mixed-selectivity to combinations of features, rather than conjunctive tuning (Fig. 3c–e; Extended Data Fig. 8k,l), and their emergence required hippocampal activity (Extended Data Fig. 8m–o).

To characterize the effect of the observed feature selectivity and mixed selectivity at the population level, we defined ensembles of neurons as feature-responsive if the mean ensemble activity to a particular feature was greater than expected by chance. We found that separate ensembles drove feature-selectivity in AC, whereas such ensembles were highly overlapping in CA1 (Fig 3f,g; Extended Data Fig. 9a, compare with Extended Data Fig. 4f). Thus, mixed-selectivity of single neurons gave rise to robust feature-selectivity at the population level, a coding strategy that has been previously highlighted in cortex^{32–34}. These AC feature ensembles, in contrast to CA1, display negligible responses to other features of the same or opposite context (Fig. 3f,g, Extended Data Fig. 9b). When AC feature-ensembles did generalize to other features, they were almost as likely to generalize to features of the opposite context, as they were to features of the same context (Fig 3h, Extended Data Fig. 9c). These findings confirm that AC exhibits a population code for features, whereas CA1 binds these features into a coherent conjunctive code of the global context.

Finally, to examine the differences in encoding mechanisms within AC and CA1, in a way that does not rely on pre-selecting feature coding ensembles, we calculated the ratio (separation index) of inter-context separation (features of opposite context) to intra-context separation (features of same context) along a hyperplane that maximally separated reward and aversive feature trials (Fig. 3i, decoding accuracy shown in Extended Data Fig. 9d, Methods). We observed that this index was closer to 1 in AC revealing higher feature separability compared to CA1 throughout the retrieval trial duration (Fig 3i; also confirmed by representing normalized mean responses of feature trials in an n-dimensional space, Extended Data Fig. 9e). Thus, AC contains robust population-level codes for features that are positioned to interact with contextual codes in CA1 during memory retrieval.

To determine whether these feature coding neurons, rather than any equally sized population of neurons in AC, are used to drive feature-based recall, we developed an approach to target

the expression of the inhibitory opsin st-GtACR preferentially in feature-coding neurons in AC. Briefly, we injected a cre-dependent st-GtACR bilaterally in AC of TRAP2 mice³⁵, and provided tamoxifen during feature-presentations (retrieval day 1) to express st-GtACR preferentially in feature coding neurons (Fig. 3j). We found, that optical silencing of these neurons (~20% of all neurons, Fig. 3j, Extended Data Fig. 9f) was sufficient to drive near complete deficit in feature-based recall (Fig. 3k left panel), which is rescued by providing no optical inhibition the following day (Fig. 3k left panel; $F(1,25)=64.39$, $p<0.0001$; Two-way ANOVA between R6 vs R7). In contrast, expression of the inhibitory opsin in a random population of neurons of a similar size (also ~20%, Extended Data Fig. 9f) and silencing during retrieval resulted in no significant deficit in feature-based recall (Fig 3k right panel; $F(1,20)=0.14$; $p>0.05$; Two-way ANOVA). Together, these results demonstrate a strong causal role for AC feature-coding neurons in driving memory recall.

Dynamic AC-CA1 reorganization enables recall

To understand how feature codes in AC emerge and interact with conjunctive codes in hippocampus during memory retrieval, we developed a dual GRIN lens preparation to perform simultaneous two-photon imaging in AC and CA1, enabled by the 2p Random Access Mesoscope³⁶ (Fig. 4a,b; Extended Data Fig. 10a,b; Supplemental Video 1, Methods). To reliably track neural sources in both regions across training and retrieval, we reduced the multi-modal contextual association task from 6 days (3 days training + 3 days retrieval) to 2 days (1 day training + 1 day retrieval). Furthermore, lengthening time in each context (60s) enabled us to track the emergence of context and feature selectivity in each region and resolve the temporal dynamics of their long range interactions (Extended Data Fig. 10c–e, Methods). We observed that during training, the recruitment of CA1 context ensembles significantly preceded the recruitment of AC ensembles, which was reversed during retrieval (Fig. 4c, $p<0.0001$ for training, $p < 0.01$ for retrieval, K–S two-tail test, Extended Data Fig. 10f). These results are consistent with a dynamic reorganization between training and retrieval where CA1 supports the emergence of AC feature representations during training (Extended Data Fig. 8m–o), which are in turn used to recruit CA1 conjunctive representations during recall (Fig. 2g, 3j,k; Extended Data Fig. 7d,e).

To further characterize this bidirectional communication between AC feature and CA1 conjunctive codes, we performed correlation analyses and, as expected, found significant correlations between the activity of feature ensembles in AC with the cognate context ensembles in CA1 (Fig. 4d, Methods). Interestingly, while AC- CA1 correlations in the neutral context were significantly above chance, these correlations were even higher for the aversive context (Fig. 4d, $q<0.05$, Friedman test), suggesting that the AC-CA1 interactions are enhanced by saliency. Indeed, most highly correlated cells across the two brain regions (Methods) consisted largely of aversive responsive neurons and had significantly more synchronous activity with aversive feature ensembles in AC and aversive context ensembles in CA1, underscoring in an unbiased way the tighter coupling of neural activity, in the aversive compared to the neutral context (Fig. 4e).

Finally, we studied how feature and context selective neurons emerge in AC and CA1 throughout training. We found that neurons that were initially shock responsive on the first two trials became context-responsive on the last two trials in both AC and CA1 (Fig. 4f,g, n=3 mice, ~60% of CA1 & ~30% of AC shock neurons became context neurons; Extended Data Fig. 10g). These data indicate that neurons initially encoding the unconditioned stimulus slowly acquire selectivity to the conditioned stimulus thus shaping a conjunctive representation of context in hippocampus. Such salient representations are further supported by bi-directional hippocampal-neocortical interactions, where hippocampus instructs salient feature representations in cortex during training, and cortex performs targeted recruitment of hippocampal representations during retrieval.

Discussion

Here we uncover a fundamental component of parallel memory processing in the brain where conjunctive representations of an experience are stored in the hippocampus CA1, while the constituent features are represented in the anterior cingulate region of the prefrontal cortex. We thus provide direct neurophysiological support to theoretical frameworks that have been developed from studies on humans^{37–39}, regarding prefrontal contributions to memory recall. We also note that feature coding in prefrontal cortex may enable organisms to recognize details in order to subsequently assign and recall associated contextual information^{9,10}, as well as identify distinct features of overlapping memory representations for high-fidelity retrieval^{24,40,41} or to extract patterns and create a semantic framework of past experiences^{42–44}, all of which are thought to be mediated by the neocortex. Moreover, dynamic access to separately stored features provides capabilities to update, modify or re-assign weights of salient features without affecting the original representation in the hippocampus. Thus, a model that aligns our findings with the literature is that parallel processing enables neocortex to parse experiences into details that are encoded in a high-dimensional manifold^{32–34,45,46} which has targeted access to hippocampal conjunctive representations encoded in a low-dimensional network to enable feature recognition and high-fidelity recall^{1,7,47}.

Furthermore, leveraging simultaneous neural circuit inhibition and imaging approaches, we find that prefrontal, but not lateral entorhinal, inputs provide targeted access to context representations in hippocampus and are required to drive memory retrieval. Notably, our observed lack of discernable feature representations in CA1 hippocampus does not contradict previous reports of unimodal representations in CA1^{3,16,17}, but rather suggests that those unimodal cues were likely considered separate contexts. Additionally, given the prominent contextual- inputs that are thought to be conveyed via lateral entorhinal cortex²⁷, we postulate that information coding may already be conjunctive by the time it reaches the entorhinal circuit^{28,29,48,49} and/or that the entorhinal-hippocampal system may primarily function during memory storage, whereas the prefrontal- hippocampal system may be a dedicated retrieval circuit (Fig. 4h).

Finally, simultaneous imaging of prefrontal cortex and hippocampus enabled us to address questions related to the co-emergence and interplay of memory representations across these regions throughout training and retrieval. We found a dynamic reorganization of temporal

structure across the hippocampal-neocortical network, with hippocampus leading cortical representations during training but a striking reversal during retrieval where top-down cortical inputs target and mobilize contextual hippocampal representations for memory recall. This bi-directional communication is enhanced by the saliency of the memory, which is likely relevant for understanding how weaker and stronger memories are represented across the brain, and how this may be altered, particularly in the prefrontal-hippocampal network. In future studies, it will be important to determine the long-term stability of the observed feature and conjunctive codes^{50–53} and whether the prefrontal-hippocampal retrieval circuit has time-limited roles, and whether molecular signatures predict functional heterogeneity that define the varying levels of commitment individual neurons have in an otherwise drifting population code.

Online Methods

Animals

All animals were purchased from The Jackson Laboratory. Six to eight weeks old wildtype C57Bl6/J male mice were group housed three to five in a cage with ad libitum food and water, unless mice were water restricted for behavioral assays. All procedures were done in accordance with guidelines approved by the Institutional Animal Care and Use Committees (protocol #19112H) at The Rockefeller University. Number of mice used for each experiment was determined based on expected variance and effect size from previous studies and no statistical method was used to predetermine sample size.

Surgeries

Viral injections—All surgical procedures and viral injections were carried out under protocols approved by the Rockefeller University IACUC. Mice were anesthetized with 1–2% isoflurane and placed in the stereotactic apparatus (Kopf) under a heating pad. Paralube vet ointment was applied on the eyes to prevent drying. Virus was injected using a 34–35 G beveled needle in a 10ul NanoFil Sub-Microliter Injection syringe (World Precision Instruments) controlled by an injection pump (Harvard Apparatus). All viruses were injected at the rate 75nl/min (unless mentioned otherwise), and the needle was removed 10mins after the injection was done to prevent backflow of the virus.

- For imaging, 500–700nl of AAV1(or AAV9)-hSyn-GCaMP6f (Addgene, Cat No. 100837-AAV1 (or AAV9); titer: $\sim 1.5 \times 10^{13}$ vg/mL) was injected in AC or CA1.
- For inhibition-imaging experiments, ~ 300 – 400 ul of a soma targeted AAV1-stGtACR2 under a CamKII promoter (Addgene, Cat No. 105669-AAV1; titer: $\sim 8 \times 10^{12}$ vg/mL) was injected unilaterally in AC/LEC.
- For optogenetic experiments, stGtACR2 was injected bilaterally in AC/LEC with the same amount and titer, with CamKII-mCherry in control cohorts (Addgene Cat No. 114469-AAV1; titer: $\sim 9 \times 10^{12}$ vg/mL).
- For chemogenetic inhibition, Gi coupled DREADD AAV9-hSyn-hM4D(Gi)-mCherry (Addgene, Cat No. 50475-AAV9; titer: $\sim 1 \times 10^{13}$ vg/mL) virus was injected in CA1 (A/P: -1.5mm, M/L: ± 1.5 mm, D/V: -1.6mm) bilaterally.

Anatomical Tracings—For retrograde tracing, 6–7 week old mice were injected with 400 nl of AAVrg-CAG-tdT (Addgene, Cat No. 59462-AAVrg; titer: $\sim 1 \times 10^{13}$ vg/mL) at 50nl/min to target dorsal hippocampus with one injection in CA1: A/P: -1.5 mm, M/L: ± 1.5 mm, D/V: -1.6 mm, and the other in CA3: A/P: -1.5 , M/L: ± 1.9 mm, D/V: -1.9 mm. Mice were housed for 4 weeks prior to perfusion and sectioning for histology. Given the enhanced sensitivity of retrograde tracers (to existing anterograde tracers), cre/flip is not necessary to easily visualize retrograde label in AC/LEC. When retroAAV-cre was used, however pre-synaptic labeling was even stronger:

- For CA1 retrograde tracing, 300nl of AAVrg-hSyn-cre (Addgene, Cat No. 105553-AAVrg; titer: $\sim 7 \times 10^{12}$ vg/mL) was injected in CA1 and 400nl of AAV1-CAG-Flex-eGFP (Addgene Cat No. 51502-AAV1; titer: $\sim 1 \times 10^{13}$ vg/mL) was injected in AC. For retrograde tracing from emx1-cre and vglut-cre mouse lines, 350 nl of AAVrg-Flex-tdT (Addgene Cat No. 28306-AAVrg; titer: $\sim 1 \times 10^{13}$ vg/mL) was injected into CA1.

For anterograde tracing, 6–7 week old mice were injected with 300nL of AAV1-hSyn-Cre (Addgene, Cat No. 105553-AAV1) in AC or LEC and 350–400nl of AAV1-CAG-Flex-eGFP (Addgene Cat No. 51502-AAV1; titer: $\sim 1.5 \times 10^{13}$ vg/mL) in CA1, and mice were housed for 4 weeks prior to perfusion and sectioning for histology.

- For dual anterograde tracing experiments, 300nL of AAV1-hSyn-Cre (Addgene, Cat No. 105553-AAV1; titer: $\sim 1.5\text{--}2 \times 10^{13}$ vg/mL) was injected in LEC, and 300nL of AAV1-Ef1a-flp (Cat No. 55637-AAV1; titer: $\sim 1 \times 10^{13}$ vg/mL) in AC, and 700nL of 1:1 mixture of AAV1-CAG-Flex-eGFP (Addgene, Cat No. 51502-AAV1; titer: $\sim 1 \times 10^{13}$ vg/mL) and AAV1-Ef1a-fDIO-mCherry (Addgene, Cat No. 114471-AAV1; titer: $\sim 1.6 \times 10^{13}$ vg/mL) in CA1. All injections were performed at 50nl/min. The experiment was also done in reverse with AAV1-hSyn-Cre injected in AC and AAV1-Ef1a-flp in LEC. Of note, using Cre/Flp substantially increased robustness and reduced technical variability for mapping these direct projections to CA1.
- To target neurons that receive convergent inputs from AC and LEC, we injected 700nL of AAV8-hSyn-ConFon-eYFP (titer: $\sim 2 \times 10^{13}$) in CA1, in conjunction with anterograde Cre and Flp systems in AC and LEC respectively. Coordinates used were; CA1 (A/P: -1.5 mm, M/L: ± 1.5 mm, D/V: -1.6 mm); AC (A/P: $+1.0$ mm, M/L: ± 0.35 mm, D/V: -1.4 mm); LEC (A/P: -4 mm, M/L: ± 3.75 mm, D/V: -4.2 mm).

Implanting coverslips, GRIN lenses, fiber optic cannulas—After confirming that mice were anesthetized (absence of reflex responses to toe-pinch), dexamethasone (0.2mg/kg) was administered intramuscularly using a 1ml syringe. Anesthesia was maintained at 1.5–2% isoflurane throughout the procedure. A long incision (covering the anteroposterior extent) was made and the skin overlying the skull was removed. The skull was cleared and dried with 3% hydrogen peroxide and scalpel scrapings.

For implanting coverslips, a 5mm craniotomy was performed with a pneumatic dental drill until a very thin layer of bone was left. This can be determined if by pushing gently on the

center of the craniotomy, the bone moves inwards. With a few drops of saline in the area, the skull was lifted with very thin forceps, keeping the dura intact. Bleeding was minimized with use of an absorbable hemostatic agent. We then positioned the circular coverslip (Harvard Apparatus CS-5R) of 5mm (thickness 1mm) over the exposed brain and sealed the remaining gap between the bone and the glass with tissue adhesive (Vetbond). A custom titanium head-plate was then placed surrounding the coverslip and glued to the skull using C&B metabond (Parkell). A 3-D printed well was glued to the head-plate using metabond and Gorilla Glue to serve as a well for holding water during two-photon imaging with a water immersion objective. We ensured a field of view primarily consisting of cingulate cortex, rather than the nearby M2 motor cortex, by targeting a field of view as close to the midline as possible, confirming that hippocampal projecting neurons were contained within our imaging field of view, and that motor related signals from these neurons were minimal and only slightly increased as we moved toward M2 and the edge of our field of view (Extended Data Fig. 8e–h).

For GRIN lens implants, a metal jacket of 1.1mm diameter was prepared from a hollow metal tube (McMaster tubing, catalog number 8987K54) and optically glued (using the Norland Optical Adhesive NOA61, Thorlabs and cured with UV light) to the 1mm diameter GRIN lens (3.4mm height, WD = 0.25mm water, Cat No. G2P10 Thorlabs) to protect the area of the lens above the skull. A craniotomy of ~1.1mm was performed and a part of cortex was suctioned out with continuous flow of PBS and low pressure vacuum. The dura beneath the craniotomy was removed using vacuum and the tip of a 28 G × 1.2” insulin syringe (Covidien). A sterile 0.5-mm burr (Fine Science Tools) attached to a stereotaxic cannula holder (Doric) was slowly inserted to the injection site and then pulled out of the brain twice to clear a pathway for the GRIN during implantation. The GRIN lens was then lowered into the brain slowly (at the rate of 1mm/min) and placed 0.2 mm above the site of injection/recording. The area surrounding the lens (metal jacket) and the skull was sealed with optic glue, and covered further with metabond. For simultaneous AC/CA1 imaging experiments, the head-plate was affixed following the small craniotomies. The head plate was placed on a forked head-bar held by a postholder on a bread board (Thorlabs) to simulate position of head-fixation during behavior. This ensured the 2 GRIN lenses were inserted at the same angle and depth, parallel to the head plate and imaging objective. We again attempted to maximize cingulate neurons in our field of view, instead of nearby M2 motor neurons, by targeting the AC GRIN lens as flush with the midline as possible, and confirming that motor related signals from our recorded neurons were minimal and increased only slight toward the lateral end of our field of view (Extended Data Fig. 10b).

For inhibition-imaging experiments, low profile cannulas were implanted in AC and LEC ipsilaterally to the recorded CA1 region, along with GRIN implants in CA1. Mono fiberoptic cannulas with low profile and NA of 0.66 (Doric, Cat No. AC: MFC_400/430–0.66_1.2mm_LPB90(P) and LEC: MFC_400/430–0.66_3.8mm_LPB90(P)_C45) allowed simultaneous imaging of CA1 via a 16x/0.8 objective without steric hindrance. The headplate was affixed to the skull first, followed by implantation of the GRIN lens and finally the low profile cannulas angled away from the center of the head. A well made of black Ortho-Jet dental cement was built around the implants using adjustable precision applicator brushes (Parkell). For optogenetics experiments, a dual fiberoptic

cannula was implanted in AC with 430um outer diameter tip in AC bilaterally with NA of 0.66 (Doric, DFC_400/430-0.66_1.1mm_GS0.7_DFL). Two mono fiberoptic cannulas were used to inhibit CA1 (Doric, MFC_400/430-0.66_1.4mm_MF1.25_DFL) and LEC (Doric, MFC_400/430-0.66_4.1mm_ZF1.25_DFL) bilaterally. All cannulas were placed 0.2mm dorsal to the site of injection, fixed to the skull with optic glue and metabond.

After surgery, animals were kept on the heating pad for recovery and given Meloxicam tablets for one day post-surgery.

Histology and Immunohistochemistry

Mice were transcardially perfused with cold PBS and 4% paraformaldehyde in 0.1M PB, then brains were post-fixed by immersion for ~24 hours in the perfusate solution followed by 30% sucrose in 0.1M PB at 4°C. Extracted brains were sliced into 40 µm coronal sections using a freezing microtome (Leica SM2010R) and stored in PBS. Free-floating sections were stained with DAPI (1:1000 in PBST), and mounted on slides with ProLong Diamond Antifade Mountant (Invitrogen). Images were acquired on a Nikon Inverted Microscope (Nikon Eclipse Ti).

For immunostaining, fixed brain sections were blocked in solution of 3% normal donkey serum, 5% BSA, and 0.2% TritonX100 in 1x PBS for ~3h and incubated with primary antibody overnight at 4°C. Sections were washed 3 times in PBS and incubated in appropriate secondary antibody for ~2.5h. Following 3x 10 min. washes in PBS, sections were stained with DAPI (1:1000 in PBST) and mounted using ProLong Diamond Antifade Mountant (Invitrogen). Images were acquired at 10X and 20X magnification with a Nikon Inverted Microscope (Nikon Eclipse Ti) and ZEISS LSM 780 confocal. Primary antibodies include rabbit monoclonal anti-cFos (Cell Signaling Technology, Cat #2250), rabbit polyclonal anti-cre recombinase (Abcam, Cat #190177), CaMKII alpha/beta/delta rabbit polyclonal antibody (Invitrogen, Cat #PA5-38239). Secondary antibodies include Alexa Fluor 647-conjugated AffiniPure donkey anti-rabbit IgG (Jackson ImmunoResearch, Cat#711-605-152, 1:250 dilution), Alexa Fluor 488 donkey anti-rabbit IgG (H+L) Highly Cross-Adsorbed secondary antibody (Thermo Fisher Scientific, Cat#A-21206). Slices were then mounted and imaged on Nikon Inverted Microscope ((Nikon Eclipse Ti).

Virtual reality behavior

We developed a virtual reality environment, adapted from previous approaches^{22,54,55}, consisting of a styrofoam ball (150mm diameter) that is axially fixed on a 10 mm rod (Thorlabs) resting on post holders, allowing motion in the forward and backward directions. Mice were head-restrained using a headplate mount above the center of the ball^{55,56}. Virtual environments were designed in ViRMEn (Virtual Reality MATLAB Engine, 2016 release) and projected onto a projector screen fabric stretched over a clear acrylic frame curved to cover ~200° of the mouse's field of view. Custom scripts in ViRMEn were used to interface the animal's behavior with the virtual environment: i.e., incoming TTL pulses from a lickometer and an optical mouse recorded animal licking and position respectively, which were used to send outgoing TTL pulses to deliver quiet solenoid-gated water/sucrose rewards, or aversive air puffs/electric shocks respectively. All sensory stimuli were provided

using a NIDAQ (NI Data Acquisition Software USB 6000, NI 782602–01) and Arduino Uno which interacted with ViRMEn MATLAB to send outgoing TTL pulses to control delivery of the various multi-modal cues, of which the olfactory and tactile cues were additionally gated by 12V quiet solenoid valves (Cat No. 2074300; The Lee Company). For auditory stimuli, a pure tone of 5 KHz, 9 KHz or 13KHz for reward, neutral and aversive contexts respectively, was provided through an Arduino speaker placed below the virtual environment; Olfactory stimuli were provided with a system of manifolds that directed air flow towards the mouse's snout. Three monomolecular odors were used, alpha-pinene (CAS Number:7785–70-8), isoamyl acetate (CAS Number :123–92-2) and methyl butyrate (CAS Number:623–42-7; Sigma Aldrich) for each of the three different contexts. Tactile stimuli consisted of controlled directional airflow (10 psi, Primefit R1401G Mini Air Regulator) behind the mouse hindpaws.

For the task, mice were water restricted (maintaining >80% of the body weight) and habituated to handling, head fixation and free water licking (~0.5 μ l per 10 licks; one 20 min session per day) for at least one week before the start of the task. During pre-exposure (Day 1) mice were repeatedly exposed, at random, to each of the three contexts (20s each), with inter-trial intervals (ITI) of 5s. During training (Days 2–4), mice again repeatedly experience each of the three contexts at random (20s ea., ensuring ~35 presentations of each context/day), together with ITIs, but this time with paired reinforcements such that the reward context was paired with sucrose delivery, the neutral context was paired with water delivery, and the aversive context was paired with aversive airpuffs. In each context, the visual cue (V) was always on, but the auditory (A), olfactory (O), and tactile (T) cues appeared intermittently to enable binding of contextual cues, where each cue was provided for a duration of 3s and appeared a total of three times during a context spanning 20s. In the reward and neutral contexts, sucrose and water were delivered continuously, whereas in the aversive context air puff was not presented continuously (to avoid habituation) but rather delivered at three distinct times. These trials were interspersed with probe trials (20%) where contexts were presented without reinforcement. Lick rate profiles (as a function of time) were calculated by averaging the lick rates across a rolling window of 1s prior of time point to 2s after time point. Integrals of these curves, averaged across trials, provided total licks per context across the session. Learning was assessed by successful modulation of lick rates on probe trials in the absence of reinforcement, i.e., enhanced licking in reward context, suppressed licking in aversive context, and no change in neutral context, with significant differences assessed by Two way ANOVA with Sidak's multiple- comparison procedure.

Once mice learned the task (significant difference in lick rate between reward and aversive context $p < 0.05$), mice move to retrieval. During retrieval (typically Days 5–7), mice were presented with 10s trials, together with 5s ITIs, consisting of full (AVOT) or partial features (AT, OT, AOT) of the original multi-modal in the absence of reinforcements (water delivery only). Full context reinforced trials were interspersed periodically (~15% of trials) to prevent extinction. We limited the types of features presented to increase number of trials and statistical power of features tested. Again, successful retrieval was assessed by appropriate modulation of lick rates in the absence of reinforcement, i.e., enhanced licking in the reward context and suppressed licking in the aversive context, with significant differences between

reward and aversive features individually assessed by a nonparametric, Mann Whitney U test.

Modification of behavior during mesoscope imaging

For simultaneous imaging of AC and CA1, we modified the behavior to observe saliency and neural dynamics on longer time scales. Thus, we presented only neutral and aversive contexts and extended the duration of each context to 60s. Scattered light from the projector screen picked up by the mesoscope PMTs (which appeared unavoidable due to the size of the objective) limited us to turn off all visual cues, and only use auditory, olfactory and tactile cues. During Training (one day only), mice were presented with 5 instances of neutral and aversive contexts of 60s each, with an ITI of 30s. Water was delivered in both contexts, but additionally, an electric tail shock (0.5mA, 0.5ms) was provided in the aversive context using the Coulbourn Precision Animal Shocker system (Harvard Apparatus; shock delivered by alligator clips attached to self-adhering conducting electrodes stuck onto the tail). In the aversive context (60s), electric shocks were delivered twice, first at 10s from context onset, and the second at a random time between 20–50s from context onset. During retrieval, each combination of partial cue was presented once (AOT, OT and AT) for 60s each, with a 15 second ITI.

Two-photon imaging during behavior

Mice were imaged throughout training (days 2–4) and retrieval (days 5–7) in ~30 min sessions/day. Volumetric imaging was performed using a resonant galvanometer two-photon imaging system (Bruker), with a laser (Insight DS+, Spectra Physics) tuned to 920 nm to excite the calcium indicator, GCaMP6f, through a 16x/0.8 water immersion objective (Nikon) interfacing with an implanted coverslip or Gradient Refractive Index (GRIN) lens through a few drops of distilled water. Fluorescence was detected through GaAs photomultiplier tubes using the Prairie View 5.4 acquisition software. Black dental cement was used to build a well around the implant to minimize light entry into the objective from the projector. High-speed z stacks were collected in the green channel (using a 520/44 bandpass filter, Semrock) at 512x512 pixels covering each x–y plane of 800x800 μm over a depth of ~150 μm (30 μm apart) by coupling the 30 Hz rapid resonant scanning (x–y) to a Z-piezo to achieve ~3.1Hz per volume. Average beam power measured at the objective during imaging sessions was between 20–40 mW. An incoming TTL pulse from ViRMEn at the start of behavior enabled time-locking of behavioral epochs to imaging frames with millisecond precision.

Simultaneous two-photon imaging of AC and CA1 during behavior

Mice were imaged throughout training and retrieval days (~15 min/session). Dual region volumetric imaging was performed using a 12KHz resonant galvanometer multiphoton mesoscope system (2p Random Access Mesoscope) with a remote focusing system for fast axial control over ~1mm range. A Tiberius Ti:Sapphire femtosecond laser (Thorlabs) was used to excite the calcium indicator GCaMP6f and fluorescence was detected using GaAsP PMTs. A water immersion objective with a 5mm aperture (0.6NA), coupled to two separate GRIN lenses was used to access both AC and CA1 in the same field of view. Remote focusing enabled rapid switching between the two axial planes corresponding to

cortical (AC) and subcortical (CA1) regions simultaneously. Volumetric images of AC and CA1 were collected over three optical planes (x/y: 600x600um), separated 60µm apart in z, achieving a volume rate of ~5.1 Hz.

Optogenetics

Mice were injected with AAV1-CaMKII-st-GTACR2-mCherry and CaMKII-mCherry (control cohort) bilaterally in AC or LEC, and implanted with dual fiber optic cannulas with guiding sockets. The cannulas (400µm diameter, 0.66NA) were placed ~0.2mm dorsal to the injection site. After recovery over 1–2 weeks, mice were water restricted, habituated to the virtual environment, and then trained on the task. During the retrieval phase, half of the trials in each trial type (AVOT, AOT, OT, AT) were inhibited (light-On), while the other half were controls (light-Off). A dual fiber optic patch-cord (DFP_400/430/2000–0.57_2m_GS0.7–2FC, Doric) was used to deliver 470nm from a laser source (DPSS Blue 473 nm Laser Cat No. MBL-III-473, Opto Engine LLC), between 4–5 mW at the tip source. Normalized lick differences between reward and aversive feature presentations was used as a measure of discrimination between reward and aversive context, referred to as the discrimination index (DI):

$$DI = \frac{(\# \text{ licks in reward feature A} - \# \text{ licks in aversive feature A})}{(\# \text{ licks in reward feature A} + \# \text{ licks in aversive feature A})}$$

Total licks from all retrieval sessions were pooled to calculate the DI. Only mice that used features to discriminate between contexts in light Off trials (i.e. DI >0.1) were included in the analysis comparing light On and light Off trials. The same cutoff was used for both st-GtACR2 and mCherry cohorts. Two-way ANOVA with Sidak's post hoc test was used to identify significant differences between DI values between *light Off* and *light On* sessions and across different feature presentations.

For CA1 inhibition, control and opsin cohorts were inhibited throughout reinforced sessions during training (T1–T3) across all contexts, and their behavior is assessed as discrimination (D.I.) between reward and aversive probe trials on last day of training (T3).

Paired Optogenetic Inhibition & Two-photon Imaging

AAV1-CamKII-GtACR was injected in AC/LEC ipsilaterally to injection of AAV-hSyn-GCaMP6f in CA1. After one week, a low-profile fiber optic cannula was implanted in AC or LEC, together with a 1mm diameter GRIN lens over CA1. Importantly, the low-profile cannulas (Doric; MFC_400/430–0.66_1.2mm_LPB90(P)_C45 for AC; MFC_400/430–0.66_3.9mm_LPB90(P)_C45 for LEC) extended outward at a 90° angle and thus did not have steric hindrance with the objective atop the GRIN lens. Mice remained in their home cages for at least one month, to allow for recovery and sufficient viral expression, before habituation to the virtual environment. Mice proceeded through training as described above. During retrieval, optogenetic inhibition of AC / LEC (on half of all trials) was paired with two-photon imaging of CA1. Inhibition of AC / LEC was achieved by activating the inhibitory opsin GtACR with 470nm light (via a Mono-fiber optic patch-cord MFP_400/430/1100–0.57_2m_FC-ZF1.25; Doric) with a maximum power of ~1–1.5mW

at the fiber tip to minimize unintended activation (and therefore increases in baseline fluorescence) of GCaMP, which would underestimate true GCaMP transients. We also confirmed that during two-photon imaging, at 920nm, with light powers less than 40mW at the objective, there was minimal unintended activation of the opsin, as evidenced by similar spontaneous and task-relevant activity patterns in the absence and presence of the opsin (Extended Data Fig. 6). Cell sources were extracted as described below (source extraction), however, in many cases the sources and portions of the resulting time series were manually verified.

Analysis of Paired Inhibition & Imaging Experiments

On inhibition trials, the 470nm light interfered only minimally with GCaMP fluorescence activity, thus introducing a small increase in baseline fluorescence, which was postprocessed after source extraction. The increase in baseline fluorescence during inhibition trials was uniform across the entire task, and thus, was corrected by performing a rolling average (~200 frames) baseline correction on the time-series from inhibited trials to match the baseline of control trials for each neuron. The responses were then z-scored in both control and inhibited trials separately and significant transients (as described below) were identified in both conditions.

CA1 neurons that were significantly inhibited by AC/LEC inhibition were defined as neurons whose mean dF activity during control trials exceeded the mean dF activity on inhibited trials by 1 standard deviation. Context selective neurons were detected from control trials (as described below) and for each condition (AC inhibition, LEC inhibition), we calculated the percentage of inhibited neurons that were also context-selective. To calculate chance probability that a context-selective neuron will be classified as inhibited, we randomly subsampled the same number of neurons as in the inhibited ensemble, from a given session, and calculated the fraction of context-selective neurons in this randomly picked subsample. This process was bootstrapped 1000x and the mean value was used as the probability by which a context-selective neuron will be inhibited by chance, rather than due to optogenetic inhibition of AC/LEC.

For calculating the relative inhibition of the activity of context-selective and other non-context selective neurons⁵⁷, separately, the time series data were binarized, where frames with significant activity (significant transients) were assigned 1, and all others set to 0. Binarizing the dF/F responses eliminated any effect of increased baseline and transient saturation emanating from the 470nm laser (if any), by avoiding magnitude comparisons and preserving the temporal structure of changes in GCaMP activity, as the measure of true response. The fraction of time a neuron in a given trial/feature is active was calculated for each neuron in both control and inhibited trials (by summing the binarized activity). The mean activation time for all context ensemble neurons in their respective feature presentations was taken as mean ensemble activity for context neurons. To account for unequal numbers of context and non-context selective neurons, we defined a “non-context” ensemble with the same number of neurons as in the context selective ensemble (non-context selective neurons were randomly picked) and the fraction of time active was calculated. This process was bootstrapped 100x and the mean of the bootstrapped values

was taken as the “mean non-context ensemble activity”. Relative inhibition was calculated as the difference of ensemble activity between control and inhibited trials normalized to the mean ensemble activity on control trials.

$$\text{Relative inhibition} = \frac{\text{Mean ensemble activity (Control)} - \text{Mean ensemble activity (Inhibited)}}{\text{Mean ensemble activity (Control)}}$$

A Two-way ANOVA with Sidak’s post hoc was performed to detect significant differences between relative inhibition of context and non-context ensembles, and mean ensemble activity differences across *light-Off* control and *light-On* trials.

c-Fos staining and overlap with AC/LEC anterograde tracing to CA1

AAV1-hSyn-Cre was injected unilaterally in AC or LEC and AAV1-CAG-Flex-eGFP injection was targeted to CA1. After 3 weeks, mice underwent Training and were perfused ~1.5 h after completing retrieval. For immediate early gene cFos staining, fixed brain sections containing CA1 were blocked in 3% normal donkey serum and 0.3% TritonX100 in 1X PBS for 1h and incubated with rabbit monoclonal anti-cFos (Cell Signaling Technology, Cat#2250, 1:200 dilution) overnight at 4°C. Sections were washed 3 times in PBS and incubated in Alexa Fluor 647-conjugated AffiniPure donkey anti-rabbit IgG (Jackson ImmunoResearch, Cat#711–605-152, 1:250 dilution) for 1.5 h, then stained with DAPI and mounted using ProLong Diamond Antifade Mountant (Invitrogen) for image collection at 10X and 20X magnification with Nikon Inverted Microscope (Nikon Eclipse Ti). The percentage of cFos+ neurons in CA1 that have overlap with AC and LEC inputs were calculated from manual cell quantification covering multiple fields of view per mouse, and quantifying across 4 mice per group. For each tracing experiment, 8 slices were used (n=4 mice, 2 slice per animal). For each animal, two slices spaced 100–150µm apart in z were quantified, where each slice had an x/y FOV spanning the entire CA1. Each data point in ED Fig 7e represents one slice.

Selective inhibition of feature coding neurons

FosTRAP transgenic mouse line TRAP2⁵⁸ were injected with AAV8-DIO-stGtACR2 in AC bilaterally and implanted with dual fiber optic cannulas with guiding sockets. The cannulas (400µm diameter, 0.66NA) were placed ~0.2mm dorsal to the injection site. After three weeks, mice were water restricted and habituated to the task setup for 5 days. On day 4 of habituation, randomly selected mice were injected with 30mg/kg of 4-hydroxy tamoxifen (4OHT), as described previously⁵⁹, to trap task irrelevant activated neurons to selectively express GtACR as controls for the experiments (TRAP habituation neurons). All mice were then trained on the task and proceeded to the retrieval phase. After day 1 of retrieval (R1), mice (excluding the control cohort of TRAP habituation neurons) were injected with 30mg/kg 4OHT to selectively express GtACR in feature responsive neurons (TRAP feature coding neurons). After 5 days from R1, AC neurons was inhibited throughout the task across all trial types for both cohorts (R6). A dual fiber optic patch- cord (DFP_400/430/2000–0.57_2m_GS0.7–2FC, Doric) was used to deliver 470nm from a laser source (DPSS Blue 473 nm Laser Cat No. MBL-III-473, Opto Engine LLC), between 4–5 mW at the tip source.

For TRAP feature coding cohort, another day of retrieval was performed the following day (R7) without any optical inhibition.

CA1 Chemogenetic inhibition and AC imaging

To chemogenetically inhibit CA1 during training, a Gi coupled DREADD (AAV9-hSyn-hM4D(Gi)-mCherry) virus was injected in CA1 (AAV9-hSyn-mCherry for controls) bilaterally along with injection of AAV9-hSyn-GCaMP6f into AC. Craniotomy and coverslip implant was performed on the mice two weeks after injection. Mice were water deprived and habituated after two weeks of recovery. 5mg/kg Clozapine N-oxide dihydrochloride (Tocris; Cat. No. 6329) was delivered intraperitoneally to both control (mCherry) and inhibited (hM4DGi) cohorts throughout training days 1–3. Retrieval was performed the following day. Mice were imaged on training day 3 and retrieval day 1.

Image Processing & Analysis

Source extraction—Two-photon images were motion-corrected using a non-rigid motion correction procedure using the NoRMCorre algorithm⁶⁰. NoRMCorre splits the field of view into overlapping spatial patches which undergoes rigid translation against a template that is continually updated. After registration, we implement a well-validated and widely used non-negative constrained matrix factorization-based algorithm, CNMF, to extract neural sources and their corresponding time series activity. Cell sources were well separated and manually assessed for any artifacts. Calcium signals were registered across days using non-rigid registration of spatial footprints using CellReg⁶¹. Calcium imaging data for dual region imaging was acquired by ScanImage 2020 software and subsequently processed using the Suite2p toolbox⁶². Motion correction, ROI detection and neuropil correction were performed as described.

Statistical Analysis of Calcium Responses—To identify statistically significant neural responses, we used an approach described Previously²² where negative going deflections in the dF/F responses are assumed to be due to motion-related artifacts and are used to estimate the fraction of positive transients that are artifactual. To do so, the number of positive and negative transients exceeding 2σ , 3σ , and 4σ over noise, were calculated, where σ was calculated on a per cell basis. The ratio of the number of negative to positive transients was calculated for different σ thresholds and for different transient durations, providing a false positive rate for each condition. We thus defined a significant transient as any transient rising above 2σ over noise for at least 2 frames, which resulted in an FPR <5%.

Identification of context- and feature- responsive (and selective) neurons—We used a probabilistic method to identify context and feature responsive neurons, both during training and retrieval. A neuron was defined as being active within a trial if its dF/F response in that trial exceeded the average response in the ITIs by 1 standard deviation. We calculated that a neuron can be active within any given trial 40% of times simply by chance, i.e., by randomly selecting intervals from the time series that match the time of a single trial, with a False Positive Rate <5% (Extended Data Fig. 2; Extended Data Fig. 8). Thus, a context responsive neuron (during training) or feature responsive neuron (during retrieval)

was defined as any neuron that was active in more than 40% of trials of that trial type. Additionally, during training, “context -selective neurons” were defined as neurons that had greater mean activity in all trials of one context compared with all trials of the other context at $p < 0.05$ by t-test with multiple comparison. During retrieval, feature-selective neurons were defined as neurons that had greater mean activity in one feature over all other features at $p < 0.05$ by one-way ANOVA.

Population trajectories—To visualize population trajectories, we began by representing the population response as an $N \times T$ matrix Y (one row for each of N neurons, one column for each of T time points, $N \ll T$)¹⁸. Entries in the matrix Y consisted of the raw activity of each neuron minus its mean over time. The columns of this matrix can be thought of as the successive coordinates of the trajectory of population activity in an N -dimensional state space, where the origin corresponds to the mean activity. To visualize these trajectories, we used singular value decomposition (SVD) (MATLAB `svd`) to find the three-dimensional subspace that captures the maximum variance in the data:

$$Y = USV^T$$

where the superscript T denotes the matrix transpose. As is standard, U is an $N \times N$ orthonormal matrix whose rows indicate how each neuron’s response is a mixture of the principal components, V is a $T \times T$ orthonormal matrix whose first N rows indicate the timecourse of each principal component (PC), and S is an $N \times T$ matrix that is nonzero only on the diagonal of its first N columns. The entries on the diagonal of S are in descending order and their squares give the variance explained by each PC. The portion of the activity explained by the first k PC’s is given by $Y_k = US_k V^T$, where S_k is the matrix S with only the first k nonzero values retained. To project Y_k onto the first k (here, $k=3$) principal axes, i.e., the 3-dimensional subspace that contains the largest amount of the variance, we computed $X = U^T Y_k = U^T US_k V^T = S_k V^T$. (The last equality follows because U is orthonormal.) The first three rows of X , as plotted in Fig. 1h, are the population trajectory in that subspace.

Decoding using a Support Vector Machine—To determine the extent to which population activity coding for context was feature dependent, we proceeded as follows. First, for each feature (OT, AT, AOT, AVOT), we trained a support vector machine (SVM) to decode context, based on all trials that contained any of the three other features, and then asked the SVM to decode the heldout trials that contained the chosen feature. The SVM decoder was trained separately for each individual time point, from 2 seconds prior to the context onset to 10s into the feature presentation. For a single time point, the model was trained on neural responses, considered as N -vectors, across trials containing any of the three non-held-out features, with the corresponding binary classification of “reward” or “aversive” label for those trials. The model was built using a linear kernel and 10-fold cross-validation. The SVM identified a hyperplane that maximally separated the neural responses from the two contexts. This hyperplane was then used to predict the context in the held-out trials, containing the responses from the held-out feature. A similar model was also built for the training phase of the task; this model was trained on reinforced trials and tested on probe trial responses to decode context.

Separation index using an optimally separating hyperplane—To understand the separation of population responses across reward and aversive contexts, and within context feature separation, we measured contrastive loss, a method to measure distances across different trial types (features) within and across contexts. First, we determined the optimally separating hyperplane for decoding context, using an SVM as described above but using all feature trials. The perpendicular to this hyperplane was then calculated, using the coefficients beta (Model.Beta in fitsvm MATLAB), and used to calculate distances between responses along the axis orthogonal to the separating hyperplane. Distance across features from the same context was termed feature separation, whereas, distance across features in opposite contexts (intercontextual) was termed context separation. Separation index was defined as the ratio of average of all inter-contextual distances (features within same context) to the average of all intra-contextual distances across (features across different context). While the separation hyperplane was able to decode between reward and aversive responses after trial onset, decoding during the ITI is at chance (50%) (Extended Data Fig. 9). Thus the “separating hyperplane” during the ITI is a random plane. We used the mean value of the separation index during the ITI to normalize the separation indices for each mouse, and then grouped data across mice and sessions.

Separation index in an N-dimensional state-space—The population response of each trial of the task was represented in an N-dimensional state space (N= no. of neurons). The mean activity of all neurons was calculated across all trial types (feature groupings) in an $N \times f$ matrix, where $f=8$ was the total number of feature groupings and contexts (AVOT, AOT, OT and AT, for both reward and aversive context). Euclidean distances in this N-dimensional state space (without the dimension reduction as described above) was calculated between positions representing features of the same context to that of features of opposite context, and the ratio of inter-contextual distance to intra-contextual distance was defined as the separation ratio (Extended Data Fig. 9).

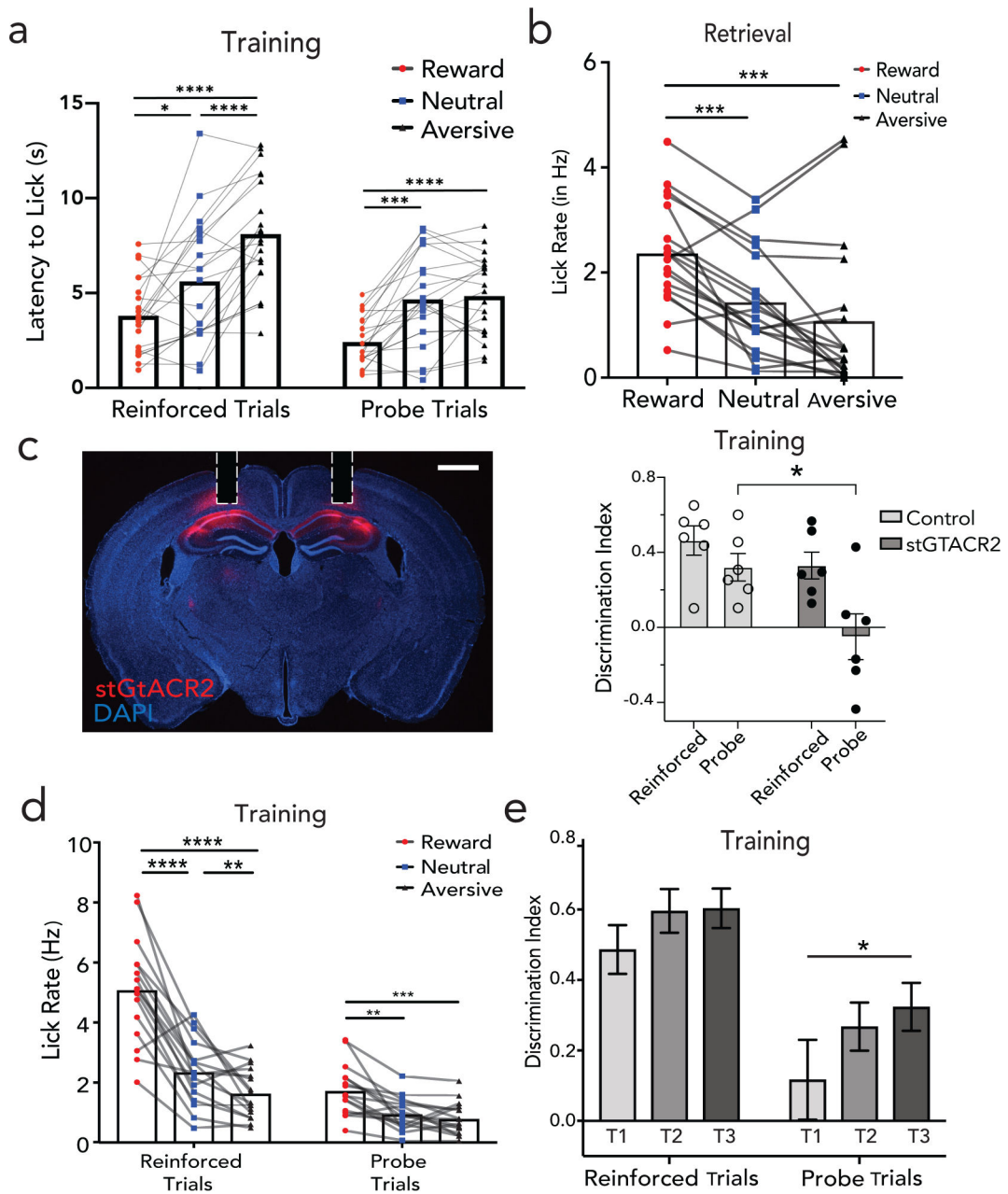
Generalized context responses—To assess the relative responses of feature ensembles to other features of the same context (generalized context responses) and features of opposite context (generalized opposite context responses), we calculated the net change in response of feature ensembles to all other features in the session. For all trial types, a net response value was calculated as the peak ensemble activity within the trial (maximum response in any 2 second window in the trial) minus the mean response in the ITI 2s prior to the trial onset. These values were then normalized to the net response of the feature ensemble in its own trial type, yielding a net relative response. The generalized context response was defined as the mean of all net relative responses to trials with features within the same context. The generalized opposite context response was defined as the mean of net relative responses to trials with features of the opposite context.

Latency analysis—To quantify the relative timing of context and feature ensembles in CA1 and AC respectively, the event onset for each neuron in these ensembles was first calculated. Event onset was defined conservatively as the first instance within a trial where the dF/F exceeded the 3σ cutoff for two consecutive frames (significant transient detection as described previously). Neurons that were not active in a given trial type were discarded

from the analysis. For the training phase of the task, the onsets for context selective neurons in both CA1 and AC were plotted as a cumulative distribution function. For the retrieval phase, context selective neurons in CA1 and feature selective neurons in AC were assessed. A Kolmogorov-Smirnov test was used to detect significant differences in the distribution of onset times.

Functionally connected neurons across regions—The functional connectedness of neurons across the two brain regions was assessed by calculating the Pearson correlation for all possible pairs across the regions. This analysis yielded a matrix of $n_a \times n_b$ correlation coefficients, where n_a is the number of neurons in region A, and n_b is number of neurons in region B. Functionally connected long-range pairs were identified as those pairs whose Pearson's correlation coefficient exceeded 0.3 (since electrophysiological studies have indicated a greater than 50% chance of in vivo functional connectedness when GCaMP correlations exceeds 0.3, thus, a conservative estimate of 0.3 was used as a cutoff)^{63,64}. Histograms of the number of functionally connected partners for each neuron were constructed, indicating the degrees distribution of the connectivity network. Highly-connected long-range neurons in both regions were defined as those neurons for which the number of correlated partners exceeded the average of the neurons in the same volume by 1 standard deviation.

Extended Data

**Extended Data Figure 1]. Behavioral performance during training and retrieval.**

a, Latency to the first lick in reward (red), neutral (blue) and aversive (black) context in both, reinforced and probe trials during training. $n=12$ mice, 24 sessions. (Two-way ANOVA with Sidak's post hoc; adjusted $*p=0.029$; $***p=0.001$; $****p<0.0001$) b, Lick rate modulation in full cue (AVOT) trials during retrieval ($n=12$ mice, 18 sessions; Two-way ANOVA with Sidak's multiple comparisons; adjusted $***p<0.005$). c, Histology of bilateral inhibition of CA1 throughout training (T1-T3, reinforced trials only, not probe trials) using st-GtACR2 and cannula implant (right, Scale: $1000\mu\text{m}$); behavioral performance measured as discrimination index (reward– aversive context lick rate / total lick rate) on probe trials

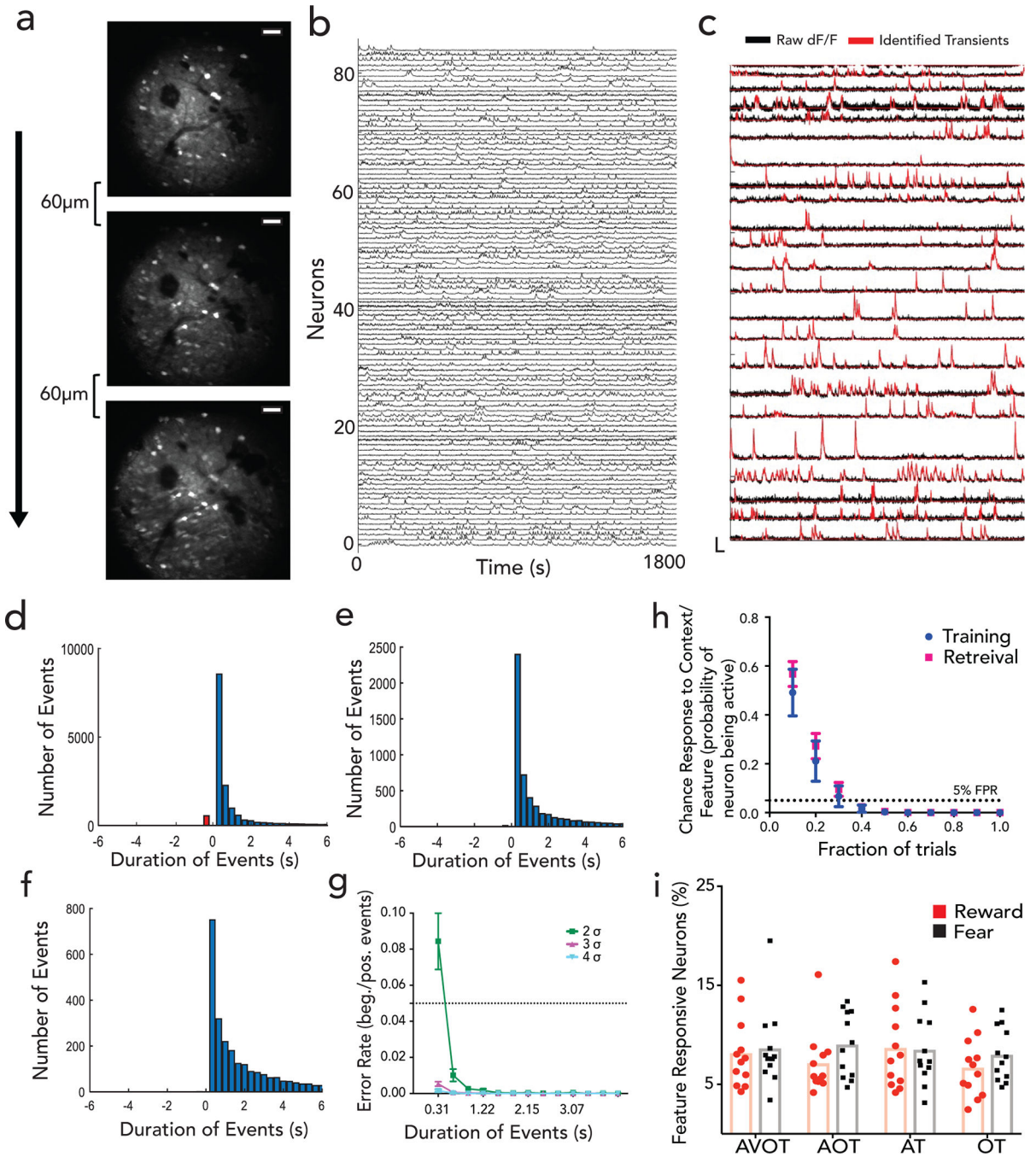
in control (mCherry) vs. opsin (GtACR) cohorts (n=6 each; Two-way ANOVA with Sidak's multiple correction, adjusted *p=0.015, Data are mean±s.e.m) d-e, GRIN lens implant does not affect learning d, lick rate modulation of mice implanted with GRINs in reinforced and probe trials, n=16 mice, (Two-way ANOVA performed with Sidak's multiple comparisons, p<adjusted *p<0.05, ***p<0.0105, ****p<0.001) e, Discrimination index in reinforced and probe trials over training days 1–3 (T1-T3); n=16 mice, Two-way ANOVA with Sidak's multiple comparison test, adjusted *p=0.01, Data are mean±s.e.m). Details of statistical analyses in Supplementary Table

Author Manuscript

Author Manuscript

Author Manuscript

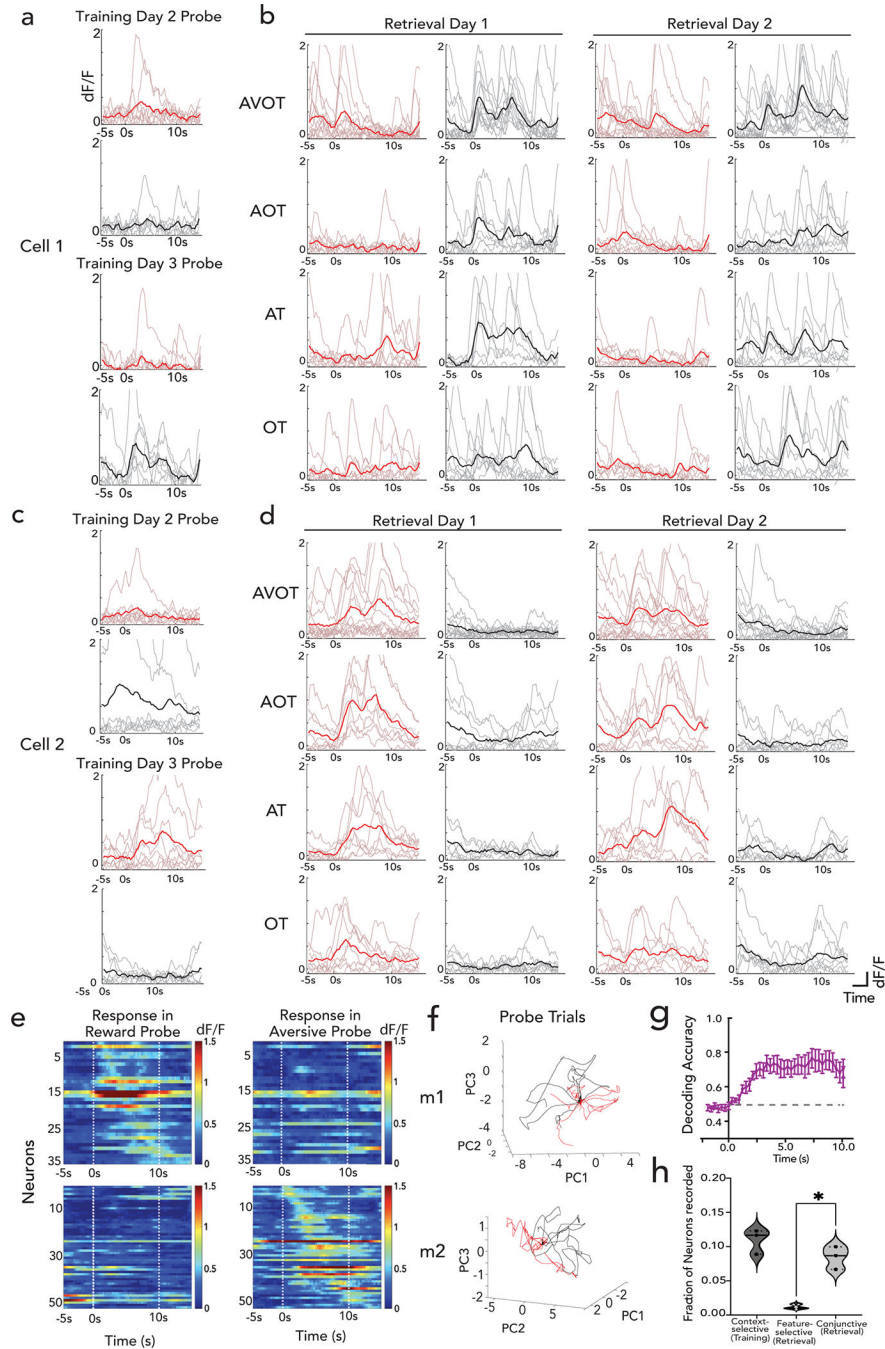
Author Manuscript



Extended Data Figure 2]. Two-photon imaging in CA1 and extraction of neural sources and activity.

a, Mean intensity Z-projections of two-photon-acquired imaging videos in CA1, showing 3 z-planes spaced 60µm apart. Scale: 50 µm, b, Example GCaMP6f neural traces during behavior with identified transients overlaid on raw dF/F activity in c. d, Identification of significant transients in dF/F traces: Histograms show distribution of positive and negative events above 2 σ thresholds over range of durations; negative going transients (red) compared to positive going transients (blue). Similar analysis for 3 σ and 4 σ threshold in e-f; g, False positive rate of transients as a function of time. FPR is described as ratio

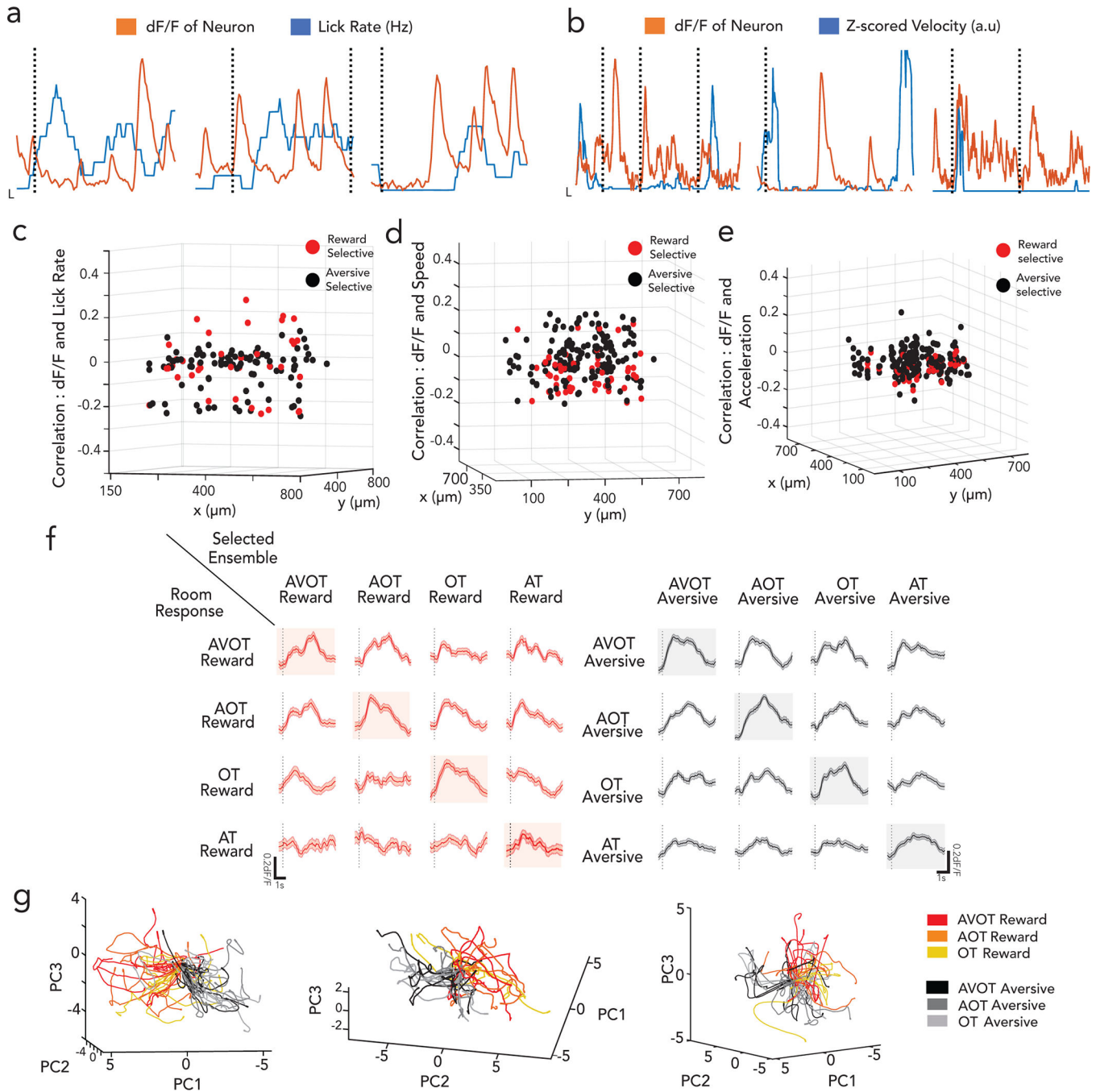
of negative to positive transients for each duration pooled across all mice (n=3 mice, 11 sessions) and fields of view and sessions (Data are mean±s.e.m). False positive curve (bottom right) is fit to an exponential curve to determine minimum transient duration and σ threshold for FPR<5%. Event onset was then described as transient going above 2 σ thresholds for 2 frames (~0.6 second) h, Chance probability of a neuron to be categorized as “feature-responsive” as a function of fraction of trials that neuron is active (n=3 mice, 9 sessions in training, 9 in retrieval). If a neuron is active on 40% of trials for a given feature, the probability of being false positive is <0.05. Data is presented as mean±sem i, %neurons classified as “feature-responsive” on retrieval trials using criteria set in h. Details of statistical analyses in Supplementary Table



Extended Data Figure 3]. Context discrimination in CA1

a-d, Single neurons registered across training days 2 and 3 (T2-T3) and retrieval days 1 and 2 (R1-R2). a, dF/F activity of neuron aligned to start of reward context probe trials (red) and aversive context probe trials (black) shows acquisition of aversive context selectivity from training day2 (top) to training day 3 (bottom), with stable context responses during aversive feature presentations in retrieval days 1–2, shown in b. c-d, same as a-b for reward context selectivity. e, Above, Heatmap shows dF/F responses of reward selective neurons in reward probe trials (left) and aversive probe trials (right) aligned to context onset at t=0

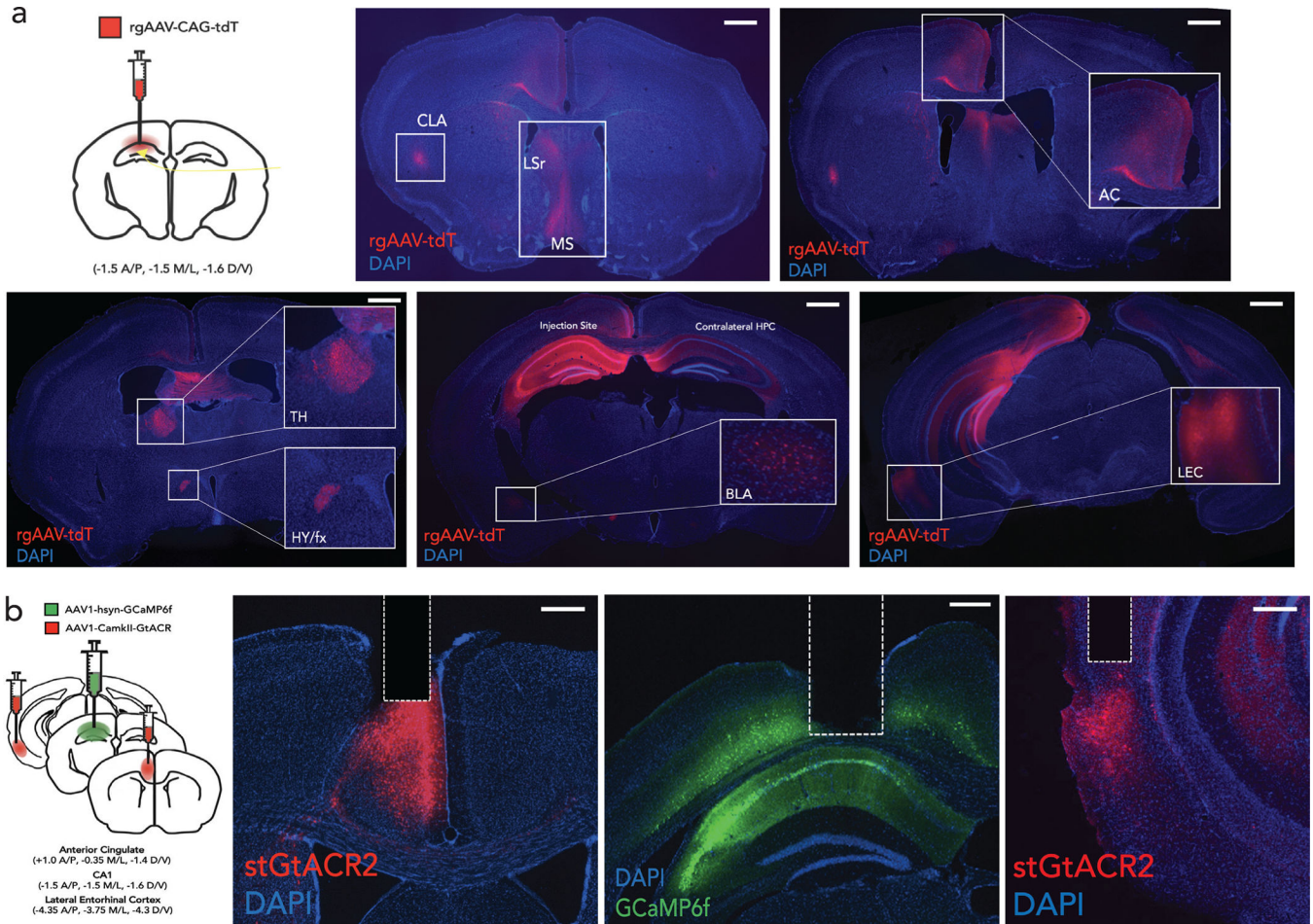
and context end at $t=10s$ (white dashed lines). Below, similar heatmap shown for aversive selective neurons in reward and aversive probe trials. f, Neural population trajectories on probe trials, similar to Fig 1h but during training, showing divergent population activity in reward (red) and aversive (black) probe trials, with variance explained by first 3 PC's (m1: 29%, m2: 19%) g, Performance of a linear SVM decoder trained on dF/F responses at each time point after context entry in reward and aversive reinforced trials and tested on probe trials shows population level discrimination between reward and aversive contexts. (n=3 mice, 6 sessions (training). Data is presented as mean \pm sem, dashed line shows chance at $y=0.5$), h, Quantification of the fraction of neurons that are context-selective (in training), feature-selective or conjunctive (in retrieval); data points represent individual mice (n=3 mice, * $p=0.024$; paired t-test). Details of statistical analyses in Supplementary Table



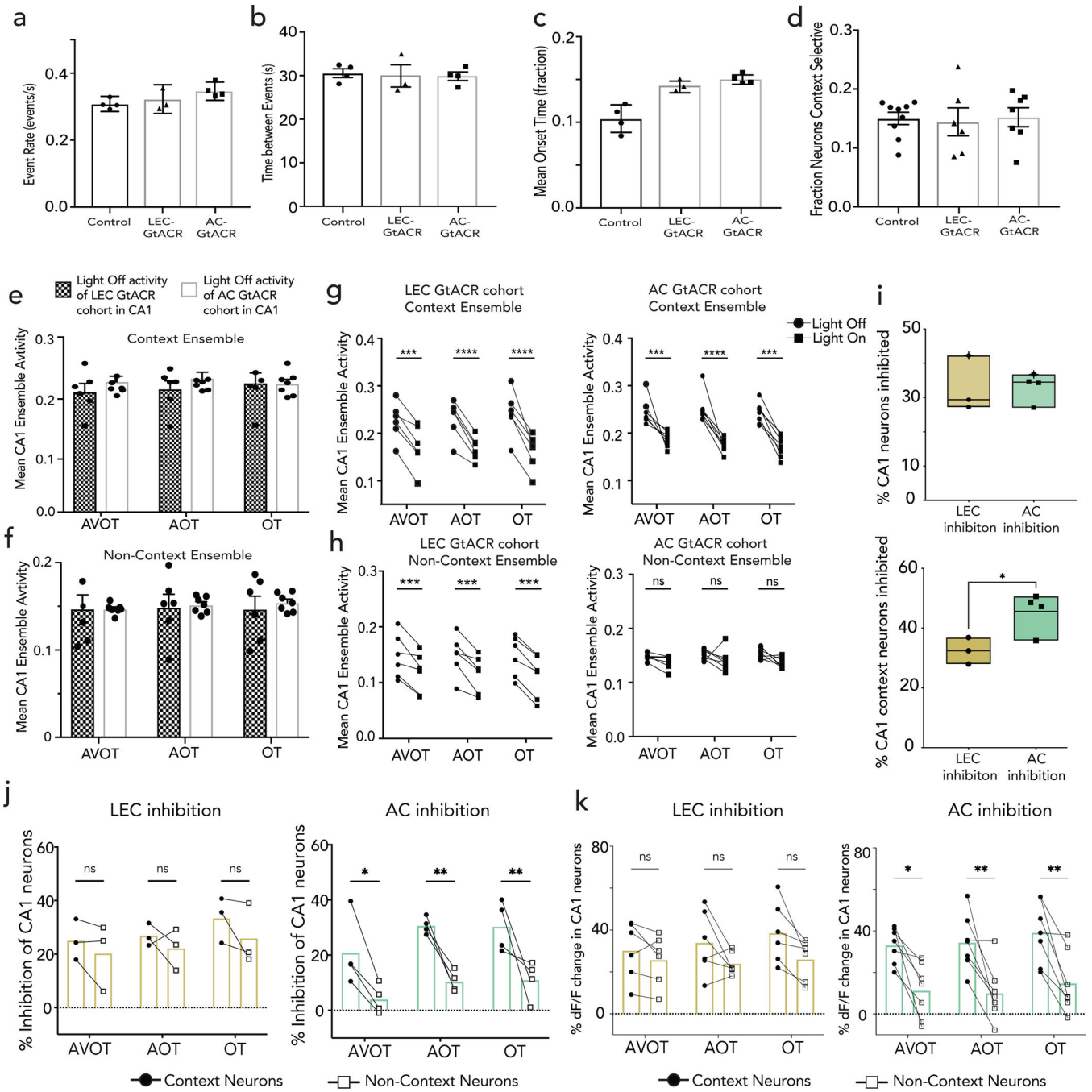
Extended Data Figure 4|. CA1 context-selective neurons are conjunctive and not movement related

a-b, sample traces of 3 context selective neurons (z-scored, orange) overlaid with lick rate (blue, in a) and speed (blue, in b), showing no significant correlation to motor activity (black dashed lines indicate onset of contexts). Context selective neurons are distributed evenly in the field of view (red: reward ensemble, black: aversive ensemble cell centroids) with low correlations (z axis) to lick rates (c), speed (d) and acceleration (e), shown here for a representative mouse. f, feature responsive ensembles exhibit highly generalized activity across all features of the same context, as shown in (g) (Scale; x:1s, y:0.2dF/F). g, Single

trial neural trajectories show indiscernible feature trajectories of the same context (similar to Fig1h but all trials projected), but divergent across the two contexts (shade of red are reward context feature trials, shades of black are aversive context feature trials) for 3 mice, with variance explained by first 3 PC's (34%, 28%, 25%).



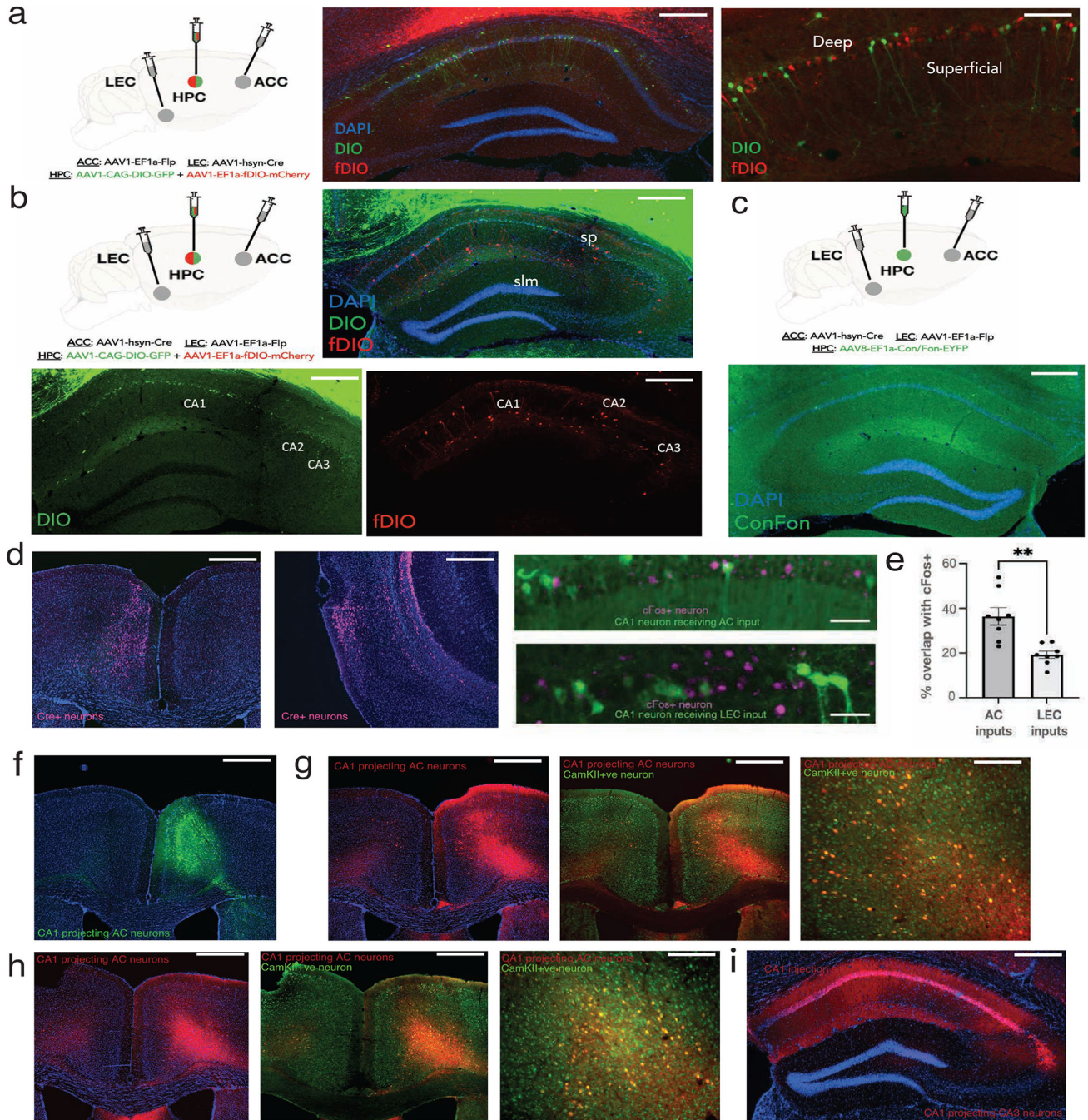
Extended Data Figure 5]. Retrograde tracing from dHPC and histology
a, Schematic of retrograde tracing from dHPC by injecting rgAAV-CAG-tdT in CA1 & CA3. Detection of retrogradely labeled tdT neurons (in red) together with nuclei stained with DAPI (blue) from coronal sections at 10x; insets show zoom at 40x; CLA: claustrum, MS: Medial septum; Lsr: Lateral Septum, AC: Anterior Cingulate, TH – Thalamus; HY/fx: Hypothalamus/fornix, BLA: Basolateral Amygdala, LEC: Lateral Entorhinal Cortex (Scale: 1000µm) b, Schematic and histology for inhibition imaging. St-GtACR2 (stained with DAPI) in AC/LEC and hSyn-GCaMP6f in CA1. Dotted line denotes placement of GRIN lens. Scale:400µm.



Extended Data Figure 6. CA1 activity in the presence and absence of opsin.

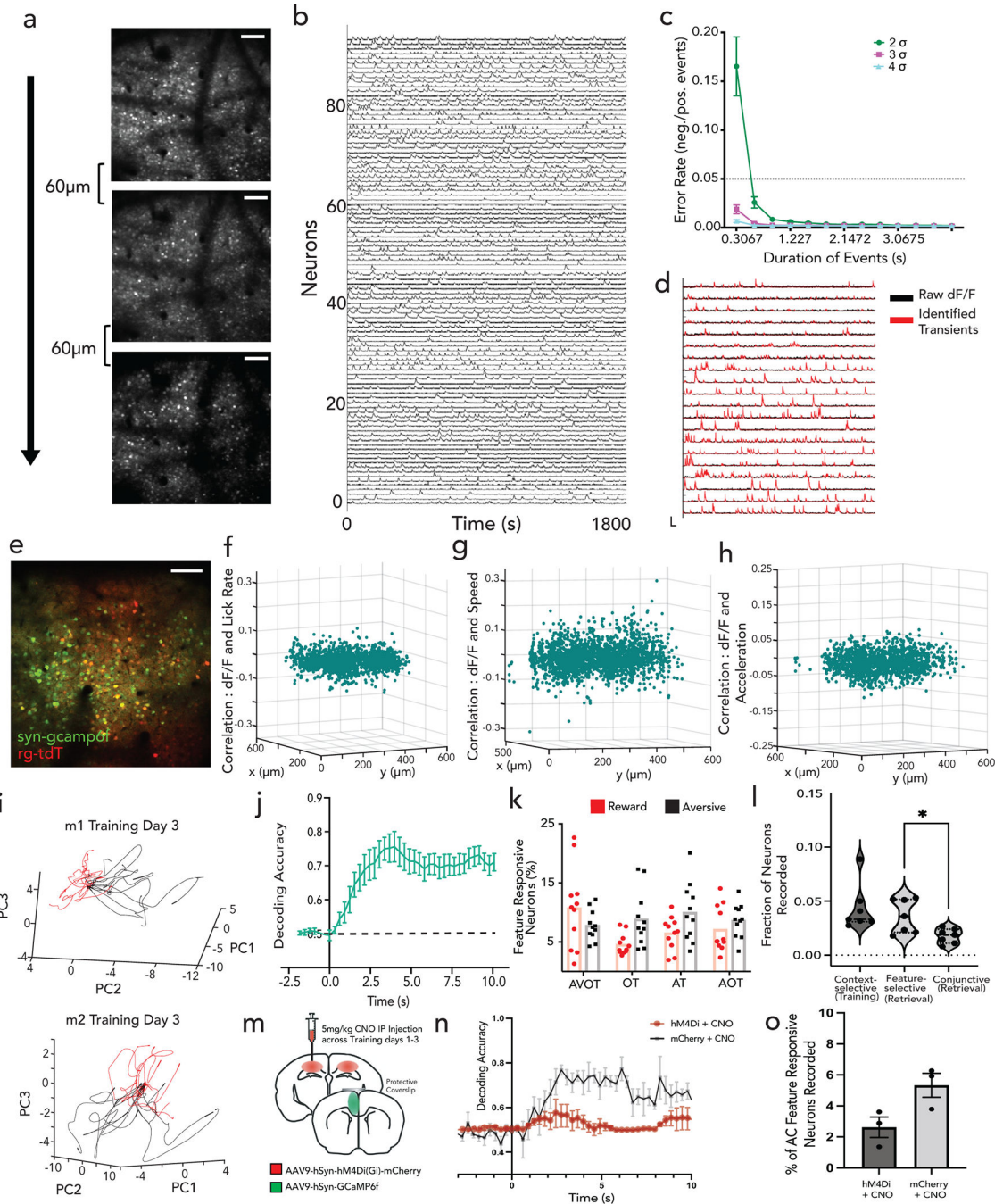
a-d, No significant activation of st-GtACR2 in AC/LEC during two-photon imaging in CA1 as assessed by similar neural activity patterns with (AC-GtACR, LEC GtACR) and without (ctrl) opsin expression. a-c, Mean event rate, time between consecutive events, and mean onset time of CA1 neurons across all three cohorts (ctrl no opsin, n=4; AC st-GtACR2, n=4; LEC st-GtACR2, n=3) with no significant differences. d, Fraction of CA1 neurons that are context selective is similar across cohorts (ctrl no opsin, n=4 mice, 9 sessions; AC st-GtACR2, n=3 mice, 7 sessions; LEC st-GtACR2, n=3 mice, 6 sessions) e-f, Mean ensemble

onset activity of CA1 neurons across both AC st-GtACR and LEC st-GtACR cohorts during light off trials (no optical inhibition) is similar across cohorts for both context-selective neurons (top) and non-context selective neurons (bottom) (All data are mean \pm s.e.m) g, Mean onset activity in CA1 during light off and light on trials in context selective ensemble (top, Two way ANOVA with Sidak's correction, adjusted p values for AVOT *** $p < 0.0001$, OT **** $p < 0.0001$) and non-context ensemble (bottom, Two way ANOVA with Sidak's correction, adjusted p values for AVOT $p = 0.008$, AOT $p = 0.0003$, OT $p = 0.0001$) for LEC inhibition cohort, signifies widespread inhibition of neurons in CA1 (n=3 mice, 6 sessions). h, same as in (g) but for AC inhibition cohort (n=4 mice, 7 sessions), context neurons (top, Two way ANOVA with Sidak's correction, adjusted p values for AVOT $p = 0.009$, AOT $p < 0.0001$, OT $p = 0.0001$) and non-context neurons (bottom, n.s.). i, Percent of all recorded CA1 (top) and CA1 context neurons (bottom) that were inhibited during AC/LEC optical inhibition (n=4 mice, AC inhibition; n=3 mice, LEC inhibition; Students t-test $p < 0.05$ for CA1 context neurons; mean, quartile, minimum and maximum are shown). j, Percent of CA1 context and non-context neurons inhibited during AC/LEC optical inhibition for all trials types (AVOT, AOT and OT) combined for reward and aversive trials, n=4 mice, AC inhibition; n=3 mice, LEC inhibition, each data point represents an individual mouse (context vs non-context neurons; for AC, $F(1,18) = 36.39$, $p < 0.001$; for LEC; $F(1,12) = 1.749$, $p = 0.21$; Two-way ANOVA with Sidak's multiple comparison) k, Percent inhibition of dF/F activity of context vs non-context neuron ensembles across all trial types (AVOT, AOT and OT) combined for aversive and reward trials, n=3 mice, 6 sessions for LEC (left), 4 mice, 7 sessions for AC (right), data are mean with each session as an individual data point; (Two-way ANOVA with Sidak's multiple comparison test; $p = 0.035$, adjusted * $p < 0.05$; ** $p < 0.01$). Details of statistical analyses in Supplementary Table.



Extended Data Figure 7 | AC and LEC project to non-overlapping subsets of neurons in CA1.
 a, Anterograde tracing from AC (AAV1-EF1a-Flp) and LEC (AAV1-hSyn-Cre) and a mixture of AAV1-CAG-GIO GFP and AAV1-EF1a-fDIO-mCherry in CA1. Neurons in red receive inputs from AC and green receive inputs from LEC, with absence of neurons having convergent input from AC and LEC (Scale: 400 μm (left), 200 μm(right)). b, Same as in (a) but injected with AAV1-hSyn-Cre in AC and AAV1-EF1a-Flp in LEC, shows lack of neurons in CA1 with convergent inputs from both AC and LEC (Scale: 400μm (left), 400μm (right)). c, (Top) Schematic of a multiplexed construct that allows a marker to be

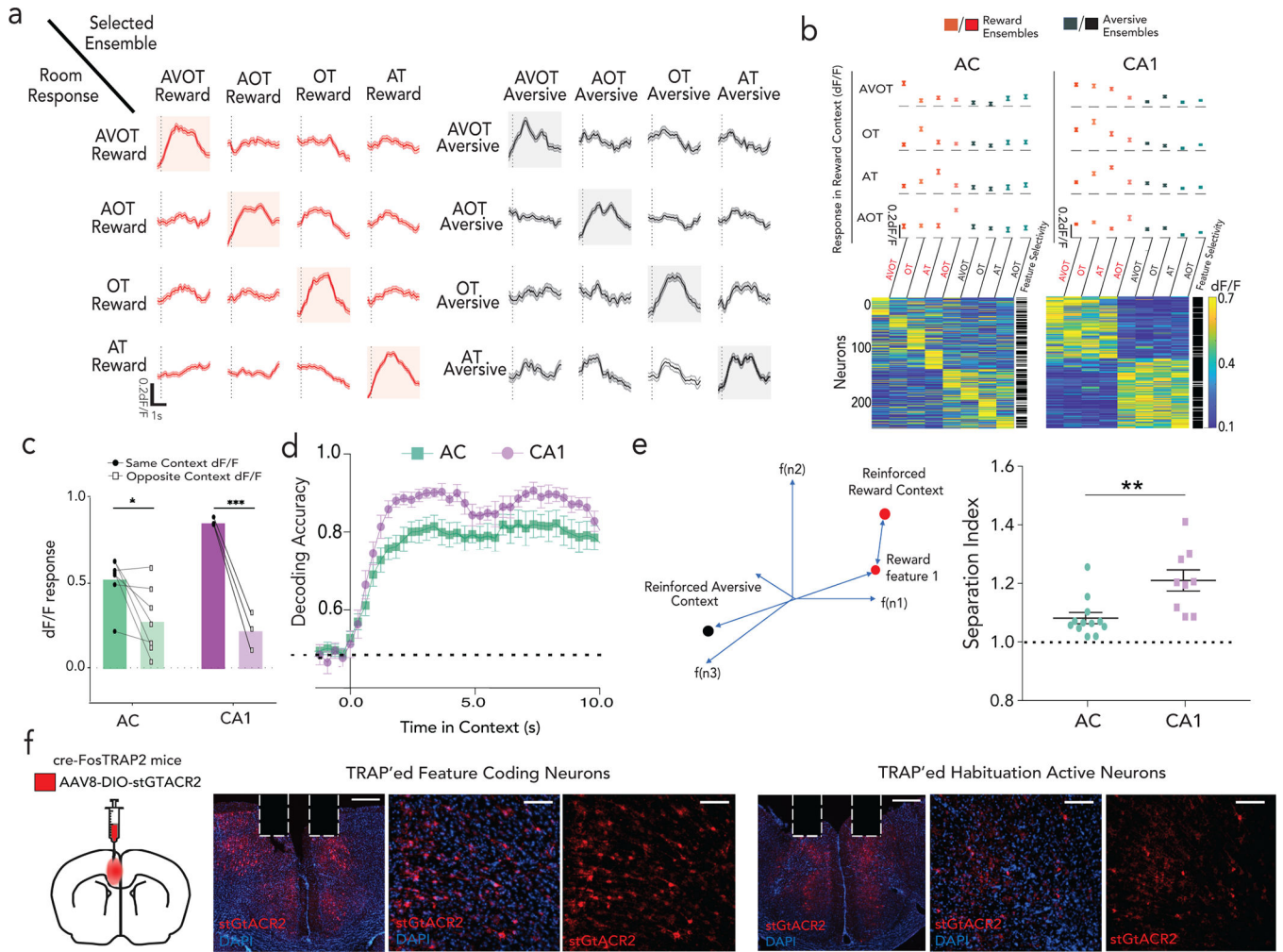
expressed only in cells that express both Cre and Flp: injecting AAV1-hSyn-Cre in AC, and AAV1-EF1a-Flp in LEC, and AAV8-EF1a-Con/Fon-eYFP in CA1. (Bottom) Coronal section of CA1 shows no neurons receive inputs from both AC and LEC. (Scale:400 μ m). d, (Left) Anterograde tracing using AAV1-hSyn-Cre, showing starter cells (stained with Cre antibody) in AC (left) or LEC (middle, overlaid with DAPI, Scale 400 μ m) with AC- (top) or LEC- receiving (bottom) CA1 neuron overlaid with cFos stain (magenta) after retrieval day 1 (right, Scale:40 μ m). e, Quantification of % of CA1 neurons receiving AC/LEC inputs that are cFos positive. LEC/AC n=4 mice, 8 slices each. p=0.002, Welch's t-test. Data points are individual slices, with mean \pm s.e.m. f, Retrograde tracing from CA1 (AAVrg-hSyn-Cre in dCA1 and AAV1-CAG-Flex-eGFP in AC) shows AC neurons (green) that project to dCA1 (Scale: 400 μ m). g-i, Retrograde tracing in emx1-cre (g) and vglut-cre (h) mice by injecting AAVrg-Flex-tdT in starter cells of dCA1(i) revealing excitatory AC-CA1 projections. CA1 projecting AC neurons (red) overlaid with DAPI (blue, left) and CaMKII stain (green, middle) (Scale: 400 μ m) and zoomed in (right, Scale: 100 μ m).



Extended Data Figure 8]. Two-photon imaging in AC during behavior

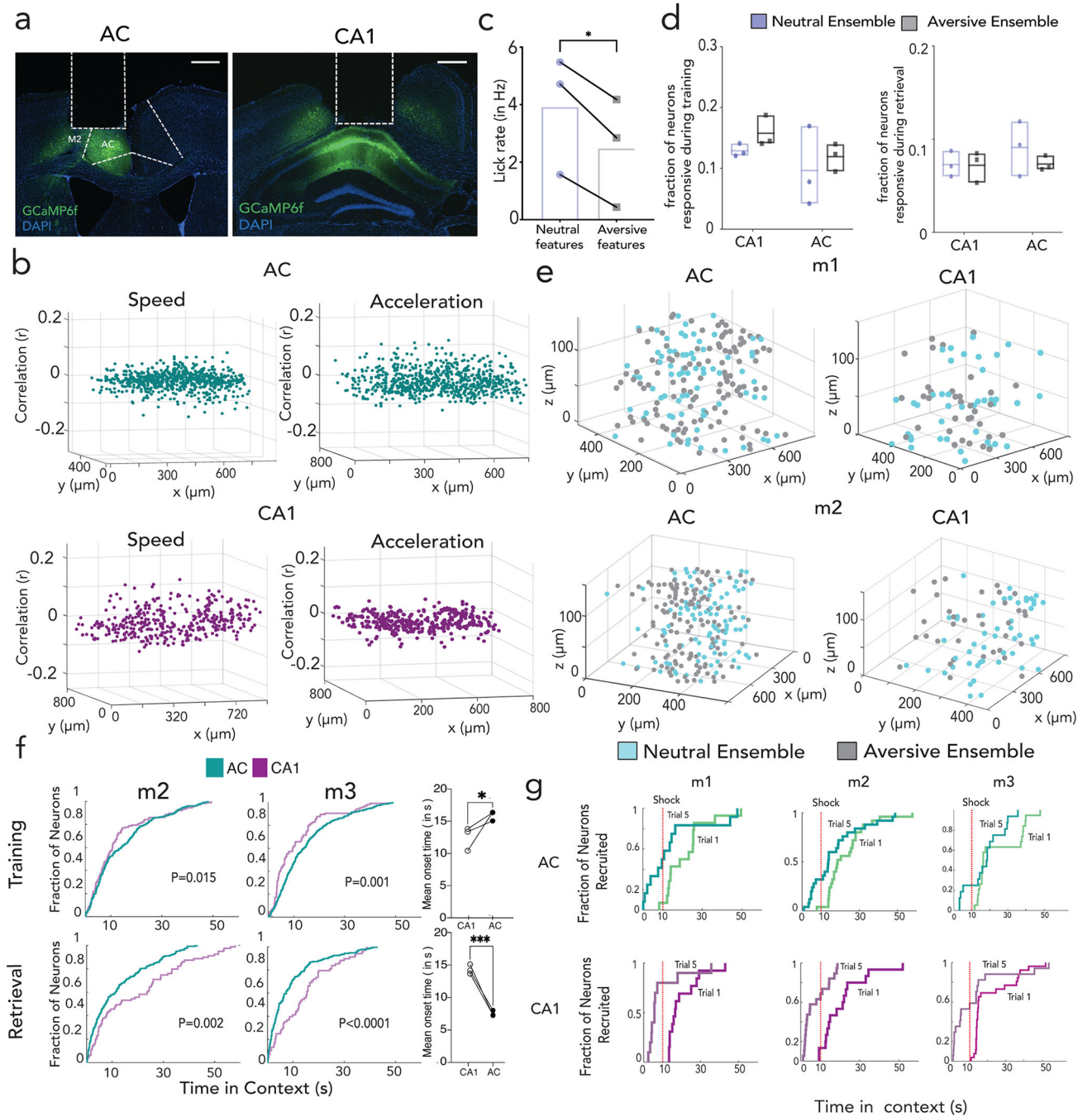
a, Mean intensity Z-projections of two-photon-acquired imaging videos in AC, showing 3 z-planes spaced 60µm apart (Scale: 50µm). **b**, Example GCaMP6f traces during behavior. **c-d**, Detecting significant transients, with false positive rate of transients as a function of time, fit to an exponential curve to determine minimum transient duration and σ threshold for FPR<5%. Event onset was then described as transient going above 2 σ thresholds for 2 frames (~0.6 second). (n=7 mice, 14 session; Data are mean±s.e.m). Detected transients overlaid with raw traces in **d**. **e**, Retrogradely labelled CA1 projecting neurons in AC (red)

in recording FOV with syn-GCaMP6f (green), indicating the FOV has direct monosynaptic access to CA1/3 (Scale: 150 μ m). **f-h**, Neuron centroids across x-y axis plotted as a function of correlations to lick rate (**f**), speed (**g**) and acceleration (**h**) shows minimal motor related signals in recording FOV. **i**, Single trial neural trajectories from representative mice (n=2 shown) show divergence between reward and aversive trajectories in probe trials, with a linear SVM decoder trained on reinforcement trials and tested on probe trials in **j**, indicating contextual discrimination in AC at population level (n=7 mice, 13 sessions in training; Data are mean \pm s.e.m). **k**, Quantification of percent of neurons responsive to each feature type across the retrieval sessions (n=7 mice, 11 sessions in retrieval). **l**, Quantification of the fraction of neurons that are context-selective (in training), feature-selective or conjunctive (in retrieval); data point represents individual mouse (n=7 mice; adjusted *p=0.015; paired t-test). **m**, Schematic of bilateral Gi-DREADD-inhibition in CA1 only across training days 1–3 while performing two-photon imaging in AC during training day 3 and retrieval, showing **n**, performance of SVM to decode context during training (n=3 mice, P<0.01 Mann Whitney U Test) and **o**, quantification of percent of AC feature responsive neurons during retrieval for CNO injected DREADD (hM4Di) vs. control (mCherry) mice (n=3 mice, n.s. paired t-test). Data are mean \pm s.e.m.



Extended Data Figure 9f. Feature ensembles in AC compared with context ensembles in CA1
a, Feature responsive ensembles in AC respond minimally to other features in the same context, shown for reward (left) and aversive (right) feature presentations. **b**, Quantification of net dF/F activity of feature responsive neurons to all other feature presentations in AC (left) and CA1 (right) from n=3 mice in reward context. Note: only context responsive neurons are shown (not all recorded neurons). Data are presented either as mean±s.e.m (above) or as heatmap (below), with adjacent column indicating whether the neuron was statistically classified as feature selective (white) or not (black). See Fig 3f for aversive **c**, Quantification of net relative response of a feature selective ensemble to other features of the same context (dark) versus the opposite context (light) in AC and CA1 respectively across all mice, each point represents individual mouse (n=7 mice ($t_{16}=2.849$; *p=0.023); n= 3 mice CA1 ($t_{16}=5.035$; ***p=0.0002); Two-way ANOVA with Sidak’s multiple comparison test). **d**, Performance of SVM to decode reward and aversive trials with all features pooled into either context (n=3 mice, 9 sessions CA1; 4 mice, 11 sessions AC; Data are mean±s.e.m). **e**, Schematic of state-space location of different features in an N-dimensional space (N is the number of neurons) and defining separation index as ratio of inter-contextual to intra-contextual distance (right). Separation index for AC (n=12 sessions) and CA1

(n= 9 sessions) (**p=0.003; Student's t-test; data as mean±s.e.m) (left). **f**, Schematic of FOSTRAP behavioral paradigm. Coronal section of AC shows neurons expressing st-GtACR2 (red) stained with DAPI (blue) with cannula implant (Scale: 500µm, 100µm). Details of statistical analyses in Supplementary Table



Extended Data Figure 10]. Simultaneous imaging of AC and CA1 during behavior.

a, Histology of AC and CA1 GRIN lens implant from a representative mouse. Scale=500µm.

b, Anatomical location of each recorded neuron and its activity correlations with motor

variables (speed, acceleration) in AC (top) and CA1 (bottom) c, Behavioral performance of dual GRIN implanted and recorded mice, shown as average lick rate on retrieval day 1 (R1) across aversive and neutral features on modified one-day behavioral paradigm (n=3 mice, *p<0.023; paired t-test). d, Left: Average fraction of context selective neurons in AC and CA1 during training day T1, with average ensemble size of feature responsive neurons in these regions during retrieval session R1 (right). e, Location of context-selective neurons are evenly distributed throughout field of view in both AC and CA1. f, Top: Proportion of context selective neurons responding (cumulative distribution function) to context onset during training for two mice as a function of latency (Kolmogorov–Smirnov two-tail test, p=0.015 (m2), p=0.001 (m3)) (purple-CA1; green-AC), with mean onset time for AC and CA1 (n=3, paired t-test p<0.05) Bottom: Same but for retrieval, with feature selective neurons in AC and context selective neurons in CA1 (Kolmogorov–Smirnov two-tail test, p=0.002(m2), p<0.0001 (m3)), with mean onset times of AC and CA1 (n=3 mice, paired t-test, p=0.01 (Training); p<0.0001 (Retrieval)). g, proportion of shock responsive neurons active after context-onset (showing first and last trials) as a function of time (cumulative distribution function), and shown separately for each individual mouse. Details of statistical analyses in Supplementary Table

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to Torsten Wiesel, Steve Siegelbaum, Ashok-Litwin Kumar, Sethu Sankaran, Vanessa Ruta, and Sachin Deshmukh for helpful discussions or critical reading of the manuscript. We thank Rani Shao in Jeff Friedman's Laboratory and Shane Johnson and Gabriela Nieves in Conor Liston's Laboratory for generously providing FosTRAP2 mice. This work is supported by a Kavli Neuroscience Institute pilot grant from the Rockefeller University, the Harold & Leila Mathers Foundation, Halis Foundation, Searle Foundation, and the National Institutes of Health under award number R00MH109652.

Data Availability :

Source data is provided with this paper. All other data that supports this study are available from the corresponding author upon request.

References

1. O'Reilly RC & Rudy JW Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review* 108, 311–345 (2001). [PubMed: 11381832]
2. O'Keefe J & Nadel L The hippocampus as a cognitive map. (Clarendon Press ; Oxford University Press, 1978).
3. Wood ER, Dudchenko PA & Eichenbaum H The global record of memory in hippocampal neuronal activity. *Nature* 397, 613–616 (1999). [PubMed: 10050854]
4. Schiller D et al. Memory and Space: Towards an Understanding of the Cognitive Map. *Journal of Neuroscience* 35, 13904–13911 (2015). [PubMed: 26468191]
5. Rolls ET The mechanisms for pattern completion and pattern separation in the hippocampus. *Front Syst Neurosci* 7, 74 (2013). [PubMed: 24198767]
6. Nakazawa K et al. Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science* 297, 211–218 (2002). [PubMed: 12040087]

7. Wills TJ, Lever C, Cacucci F, Burgess N & O'Keefe J Attractor dynamics in the hippocampal representation of the local environment. *Science* 308, 873–876 (2005). [PubMed: 15879220]
8. Leutgeb JK, Leutgeb S, Moser M-B & Moser EI Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315, 961–966 (2007). [PubMed: 17303747]
9. Kirwan CB, Wixted JT & Squire LR Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. *J Neurosci* 28, 10541–10548 (2008). [PubMed: 18923030]
10. Eichenbaum H, Yonelinas AP & Ranganath C The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30, 123–152 (2007). [PubMed: 17417939]
11. Marr D Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci* 262, 23–81 (1971). [PubMed: 4399412]
12. Alberini CM Mechanisms of memory stabilization: are consolidation and reconsolidation similar or distinct processes? *Trends Neurosci* 28, 51–56 (2005). [PubMed: 15626497]
13. Schwindel CD & McNaughton BL Hippocampal–cortical interactions and the dynamics of memory trace reactivation. in *Progress in Brain Research* vol. 193 163–177 (Elsevier, 2011). [PubMed: 21854962]
14. Duncan J An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci* 2, 820–829 (2001). [PubMed: 11715058]
15. Roy JE, Riesenhuber M, Poggio T & Miller EK Prefrontal Cortex Activity during Flexible Categorization. *Journal of Neuroscience* 30, 8519–8528 (2010). [PubMed: 20573899]
16. Eichenbaum H, Kuperstein M, Fagan A & Nagode J Cue-sampling and goal-approach correlates of hippocampal unit activity in rats performing an odor-discrimination task. *J Neurosci* 7, 716–732 (1987). [PubMed: 3559709]
17. Kaifosh P & Losonczy A Mnemonic Functions for Nonlinear Dendritic Integration in Hippocampal Pyramidal Circuits. *Neuron* 90, 622–634 (2016). [PubMed: 27146266]
18. Freeman J et al. Mapping brain activity at scale with cluster computing. *Nat Methods* 11, 941–950 (2014). [PubMed: 25068736]
19. Fyhn M, Hafting T, Treves A, Moser M-B & Moser EI Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446, 190–194 (2007). [PubMed: 17322902]
20. Deshmukh SS & Knierim JJ Influence of local objects on hippocampal representations: Landmark vectors and memory. *Hippocampus* 23, 253–267 (2013). [PubMed: 23447419]
21. Komorowski RW, Manns JR & Eichenbaum H Robust Conjunctive Item-Place Coding by Hippocampal Neurons Parallels Learning What Happens Where. *Journal of Neuroscience* 29, 9918–9929 (2009). [PubMed: 19657042]
22. Rajasethupathy P et al. Projections from neocortex mediate top-down control of memory retrieval. *Nature* 526, 653–659 (2015). [PubMed: 26436451]
23. Tomita H, Ohbayashi M, Nakahara K, Hasegawa I & Miyashita Y Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* 401, 699–703 (1999). [PubMed: 10537108]
24. Xu W & Sudhof TC A Neural Circuit for Memory Specificity and Generalization. *Science* 339, 1290–1295 (2013). [PubMed: 23493706]
25. Ito HT, Zhang S-J, Witter MP, Moser EI & Moser M-B A prefrontal–thalamo–hippocampal circuit for goal-directed spatial navigation. *Nature* 522, 50–55 (2015). [PubMed: 26017312]
26. Malik R, Li Y, Schamiloğlu S, Sohal VS. Top-down control of hippocampal signal-to-noise by prefrontal long-range inhibition. *Cell*, Volume 185, Issue 9, 2022
27. Hargreaves EL, Rao G, Lee I & Knierim JJ Major dissociation between medial and lateral entorhinal input to dorsal hippocampus. *Science* 308, 1792–1794 (2005). [PubMed: 15961670]
28. Deshmukh SS & Knierim JJ Representation of Non-Spatial and Spatial Information in the Lateral Entorhinal Cortex. *Front. Behav. Neurosci.* 5, (2011).
29. Basu J, Zaremba JD, Cheung SK, Hitti FL, Zemelman BV, Losonczy A, Siegelbaum SA Gating of hippocampal activity, plasticity, and memory by entorhinal cortex long-range inhibition. *Science* 351, aaa5694–aaa5694 (2016). [PubMed: 26744409]

30. Woods NI et al. The Dentate Gyrus Classifies Cortical Representations of Learned Stimuli. *Neuron* 107, 173–184.e6 (2020). [PubMed: 32359400]
31. Mahn M et al. High-efficiency optogenetic silencing with soma-targeted anion-conducting channelrhodopsins. *Nat Commun* 9, 4125 (2018). [PubMed: 30297821]
32. Rigotti M et al. The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590 (2013). [PubMed: 23685452]
33. Mante V, Sussillo D, Shenoy KV & Newsome WT Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78–84 (2013). [PubMed: 24201281]
34. Raposo D, Kaufman MT & Churchland AK A category-free neural population supports evolving demands during decision-making. *Nat Neurosci* 17, 1784–1792 (2014). [PubMed: 25383902]
35. DeNardo LA, Liu CD, Allen WE, et al. Temporal evolution of cortical ensembles promoting remote memory retrieval. *Nat Neurosci*. 22(3):460–469 (2019). [PubMed: 30692687]
36. Sofroniew NJ, Flickinger D, King J & Svoboda K A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *eLife* 5, e14472 (2016). [PubMed: 27300105]
37. Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*. 1998 Aug 21;281(5380):1188–91. [PubMed: 9712582]
38. Miller EK, & Cohen JD (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
39. Koechlin E, Ody C, Kouneiher F. The architecture of cognitive control in the human prefrontal cortex. *Science*. 2003 Nov 14;302(5648):1181–5. [PubMed: 14615530]
40. Guise KG & Shapiro ML Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. *Neuron* 94, 183–192.e8 (2017). [PubMed: 28343868]
41. Jo YS et al. The Medial Prefrontal Cortex Is Involved in Spatial Memory Retrieval under Partial-Cue Conditions. *Journal of Neuroscience* 27, 13567–13578 (2007). [PubMed: 18057214]
42. Yu JY, Liu DF, Loback A, Grossrubatscher I & Frank LM Specific hippocampal representations are linked to generalized cortical representations in memory. *Nat Commun* 9, 2209 (2018). [PubMed: 29880860]
43. Bernardi S et al. The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell* 183, 954–967.e21 (2020). [PubMed: 33058757]
44. Reinert S, Hübener M, Bonhoeffer T & Goltstein PM Mouse prefrontal cortex represents learned rules for categorization. *Nature* (2021) doi:10.1038/s41586-021-03452-z.
45. Beck JM et al. Probabilistic population codes for Bayesian decision making. *Neuron* 60, 1142–1152 (2008). [PubMed: 19109917]
46. Stringer C, Pachitariu M, Steinmetz N, Carandini M & Harris KD High-dimensional geometry of population responses in visual cortex. *Nature* 571, 361–365 (2019). [PubMed: 31243367]
47. Yoon K et al. Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nat Neurosci* 16, 1077–1084 (2013). [PubMed: 23852111]
48. Zhao X, Wang Y, Spruston N & Magee JC Membrane potential dynamics underlying context-dependent sensory responses in the hippocampus. *Nat Neurosci* 23, 881–891 (2020). [PubMed: 32451487]
49. Sargolini F Conjunctive Representation of Position, Direction, and Velocity in Entorhinal Cortex. *Science* 312, 758–762 (2006). [PubMed: 16675704]
50. Ziv Y et al. Long-term dynamics of CA1 hippocampal place codes. *Nat Neurosci* 16, 264–266 (2013). [PubMed: 23396101]
51. Driscoll LN, Pettit NL, Minderer M, Chettih SN & Harvey CD Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. *Cell* 170, 986–999.e16 (2017). [PubMed: 28823559]
52. Richards B & Frankland P (2017). The Persistence and Transience of Memory. *Neuron*. 94. 1071–1084. [PubMed: 28641107]
53. Kitamura T, Ogawa SK, Roy DS, Okuyama T, Morrissey MD, Smith LM, Redondo RL, Tonegawa S. Engrams and circuits crucial for systems consolidation of a memory. *Science*. 2017 Apr 7;356(6333):

Methods References

54. Dombeck DA, Harvey CD, Tian L, Looger LL & Tank DW Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat Neurosci* 13, 1433–1440 (2010). [PubMed: 20890294]
55. Harvey CD, Coen P & Tank DW Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* 484, 62–68 (2012). [PubMed: 22419153]
56. Niell CM & Stryker MP Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* 65, 472–479 (2010). [PubMed: 20188652]
57. Chettih SN and Harvey CD Single-neuron perturbations reveal feature-specific competition in V1. *Nature* 567, 334–340 (2019). [PubMed: 30842660]
58. Guenther CJ, Miyamichi K, Yang HH, Heller HC, Luo L. Permanent genetic access to transiently active neurons via TRAP: targeted recombination in active populations. *Neuron*. 2013 Jun 5;78(5):773–84 [PubMed: 23764283]
59. Ye L, Allen WE, Thompson KR, et al. Wiring and Molecular Features of Prefrontal Ensembles Representing Distinct Experiences. *Cell*. 2016;165(7):1776–1788. doi:10.1016/j.cell.2016.05.010 [PubMed: 27238022]
60. Pnevmatikakis EA & Giovannucci A NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods* 291,83–94 (2017). [PubMed: 28782629]
61. Sheintuch L et al. Tracking the Same Neurons across Multiple Days in Ca²⁺ Imaging Data. *Cell Rep* 21, 1102–1115 (2017). [PubMed: 29069591]
62. Pachitariu M et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. 10.1101/061507 (2016) doi:10.1101/061507.
63. Cossell L et al. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* 518, 399–403 (2015). [PubMed: 25652823]
64. Ko H. et al. . Functional specificity of local synaptic connections in neocortical networks. *Nature* 473, 87–91 (2011). [PubMed: 21478872]

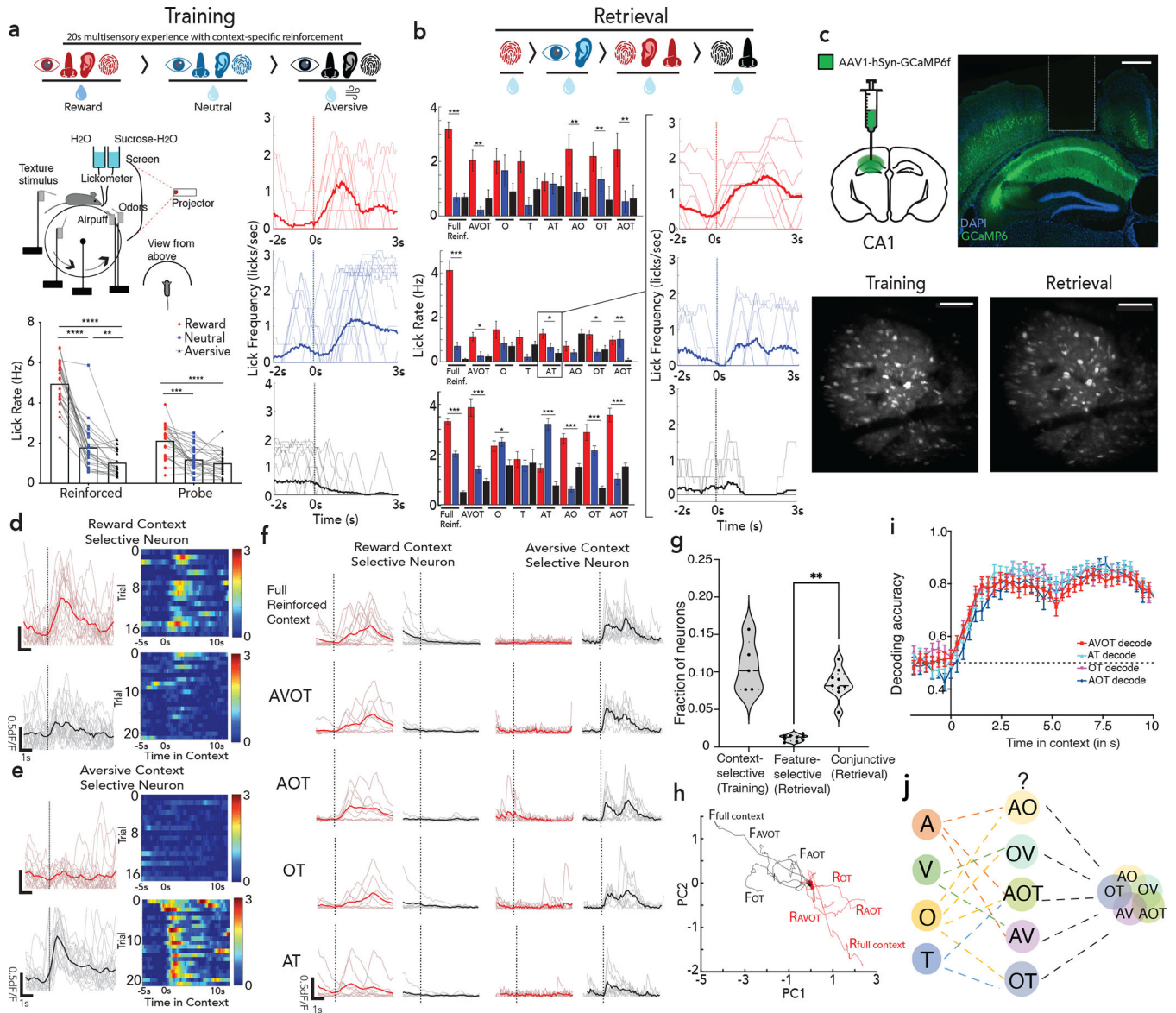


Figure 1 |. The hippocampus CA1 encodes conjunctive representations of context

a, Head fixed virtual reality setup with training paradigm (top); Lick rates in reward, neutral and aversive context across reinforced and probe trials, n=12 mice, 24 sessions (**p<0.01, ****p<0.0001, Two-way ANOVA with Sidak’s post-hoc, bottom left); with moving average lick rates across probe trials (translucent lines) from one mouse aligned to context entry (bottom right, session average in opaque lines). **b**, Retrieval paradigm; A – auditory, V – Visual, O – olfactory, T – texture (top). Lick rate on full (AVOT) or partial features (O, T, AT, AO, OT, AOT) of the respective contexts for 3 mice (mean±s.e.m, *p<0.05, **p<0.01 Two-way ANOVA with multiple comparison); moving average lick responses from representative mouse (right). **c**, Histology (scale: 500um) and Z-projection images of two-photon recording (mean over time, scale:160um) of GCaMP expressing dorsal CA1 neurons with GRIN implants capturing same field of view (FOV) across training and retrieval **d-e**, Activity of a reward (**d**) and aversive (**e**) context-selective neuron in reward

(red) and aversive (black) probe trials (mean response in opaque line) with heatmap showing individual trial responses. **f**, Representative neurons exhibiting conjunctive representation of reward (right) or aversive (left) contexts showing significant responses to all features of only one context. **g**, Fraction of context-selective (in training), feature-selective or conjunctive neurons (in retrieval) (n=3 mice, 5 sessions for training, 9 sessions in retrieval; paired t test during retrieval, $**p=0.003$). **h**, Significant divergence on retrieval trials of reward (red) vs aversive (black) population trajectories from a representative mouse, **i**, Performance of a linear SVM to classify context, trained on three features and tested on a held-out feature of the same context, demonstrating shared underlying dynamics for all features of a given context (n=3 mice, 9 sessions, mean \pm s.e.m). **j**, schematic posing the question: where are features represented and how do they access CA1 conjunctive representations. Details of statistical analyses in Supplementary Table

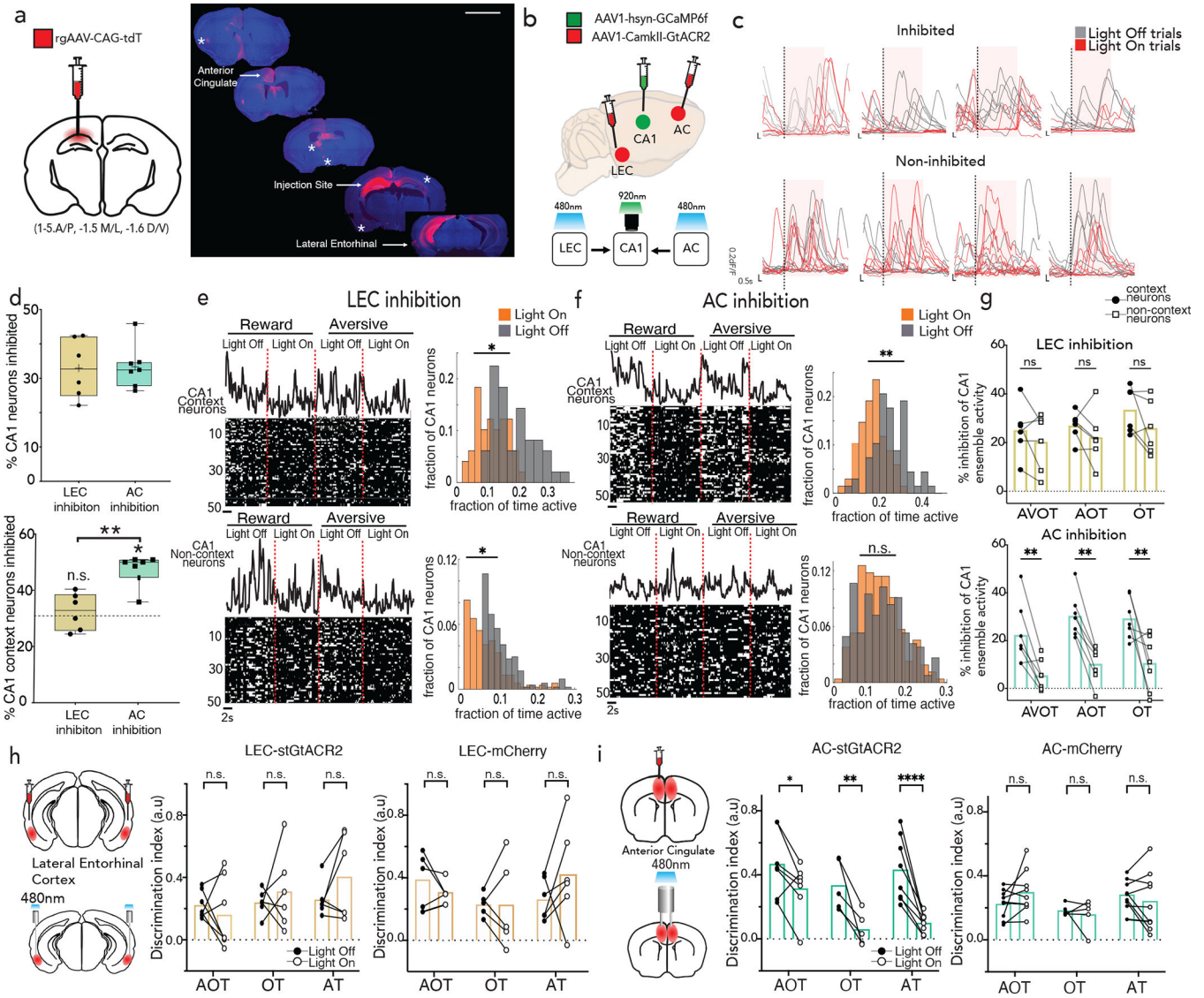


Figure 2 | AC, not LEC, is necessary for feature based memory recall.

a, RetroAAV-tdT injection in dorsal hippocampus labeled afferents from anterior cingulate (AC), lateral entorhinal cortex (LEC), claustrum, medial septum, anterior thalamus, hypothalamus, contralateral hippocampus, and amygdala (denoted with asterisks) (scale: 3000um, see Extended Data Fig. 7c–e). **b**, AC/LEC inhibition with simultaneous CA1 imaging, shows **c**, robust inhibition of some neurons (top: light on inhibition trials in red, control light off trials in black), and not others (bottom). Shaded area represents light delivery (Scale: y:0.2 dF/F; x:0.5s) **d**, No significant difference between fraction of CA1 neurons inhibited during LEC (n=4 mice, 7 sessions) vs. AC (n=3 mice, 6 sessions) inhibition (top) but CA1 context neurons are preferentially inhibited during AC inhibition compared to LEC (bottom, Student’s t-test *p=0.001) and by chance (dashed line represents chance, Wilcoxon signed-rank test *p=0.015) (mean, quartile, minimum and maximum are shown). **e,f**, Rasters of binarized activity from 50 neurons in CA1 during LEC (**e**) and AC inhibition (**f**) in reward and aversive feature trials grouped by context responsive vs

non-context neurons from one mouse; (right) activity of neurons quantified in histograms. Significant left shift (* $p < 0.05$, Wilcoxon signed-rank test) observed on inhibited trials across both groups during LEC inhibition, but only of context specific neurons during AC inhibition (** $p < 0.01$). **g**, Percent inhibition of context vs non-context neuron activity across all trial types (AVOT, AOT and OT) combined for aversive and reward trials, $n = 3$ mice, 6 sessions for LEC (top), 4 mice, 7 sessions for AC (bottom), each session is an individual data point ($F_{(1,36)} = 38.92$, $p < 0.0001$ for AC; $F_{(1,30)} = 2.801$, $p = 0.11$ for LEC; Two-way ANOVA with Sidak's post-hoc) **h-i**, Bilateral optogenetic-inhibition of LEC(h) or AC(i) during memory retrieval showing behavioral performance during inhibited (*light on*) vs control (*light off*) trials for each trial type, in opsin (GtACR) vs. control (mCherry) cohorts. Data points are individual mice ($F_{(1,17)} = 43.79$, $p < 0.0001$ for AC GtACR; $F_{(1,16)} = 0.82$, $p = 0.37$ for LEC GtACR cohort; Two-way ANOVA with Sidak's post-hoc). Details of statistical analyses in Supplementary Table

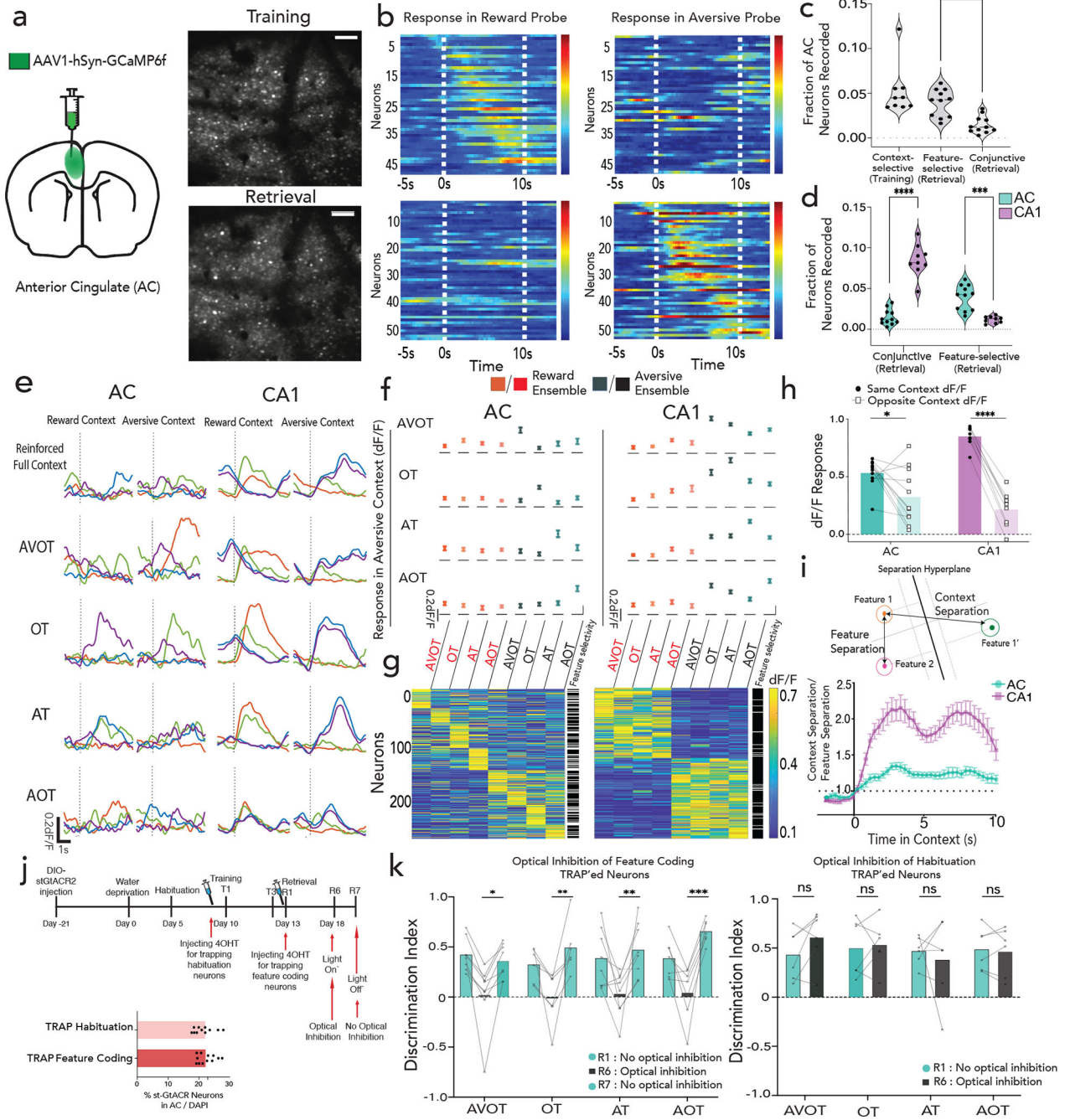


Figure 3 | Population codes for feature representation in AC

a, Schematic and Z-projection images (scale:160um) of two-photon recording in AC. **b**, Reward (above) and aversive (below) context-selective neurons, sorted by onset time, displayed for reward (left) and aversive (right) probe trials from one mouse. **c**, Fraction of context-selective (in training), feature-selective or conjunctive (in retrieval) neurons in AC (n=7 mice, 9 sessions in training, 11 sessions in retrieval; paired t-test in retrieval $**p=0.002$); with a comparison of conjunctive and feature responses in AC and CA1 in (**d**) (n=3 mice, 9 sessions CA1, n=7 mice, 11 sessions AC, $****p<0.0001$, $**p=0.0005$;

Two-way ANOVA with Sidak's post-hoc). **e**, Representative examples of four neurons, each in a different color, exhibiting feature selectivity in AC and conjunctive representation in CA1 during retrieval. **f,g**, dF/F activity of feature responsive neurons to all other feature presentations in AC (left) and CA1 (right, n=3 mice each), presented as either as mean±s.e.m (f) or heatmap (g), with adjacent column indicating whether the neuron was statistically classified as feature selective (white) or not (black). **h**, Relative dF/F response of a feature selective ensemble to other features of the same context (dark) versus the opposite context (light) in AC and CA1 (n=7 mice, 11 sessions AC; n= 3 mice, 9 sessions CA1; multiple paired t-test ****q<0.0001; *q=0.027; see Extended Fig 9c). **i**, Ratio of context to feature separation across a maximally separating hyperplane in state space on retrieval trials (n=7 mice, 11 sessions AC; n= 3 mice, 9 sessions CA1, Data are mean±s.e.m). **j**, Timeline of TRAP2 paradigm to inhibit feature coding neurons (top), with % AC neurons that expressed GtACR in the habituation TRAP'ed and feature coding TRAP'ed cohort **k**, TRAP'ed mice display deficits in feature-based recall on R6, which is rescued during *light-off* on R7 ($F_{(1,25)}=64.39$, $p<0.0001$). No behavioral deficit in habituation TRAP'ed cohort ($F_{(1,20)}=0.14$; $p=0.71$; Two-way ANOVA with Sidak's post-hoc). Details of statistical analyses in Supplementary Table.

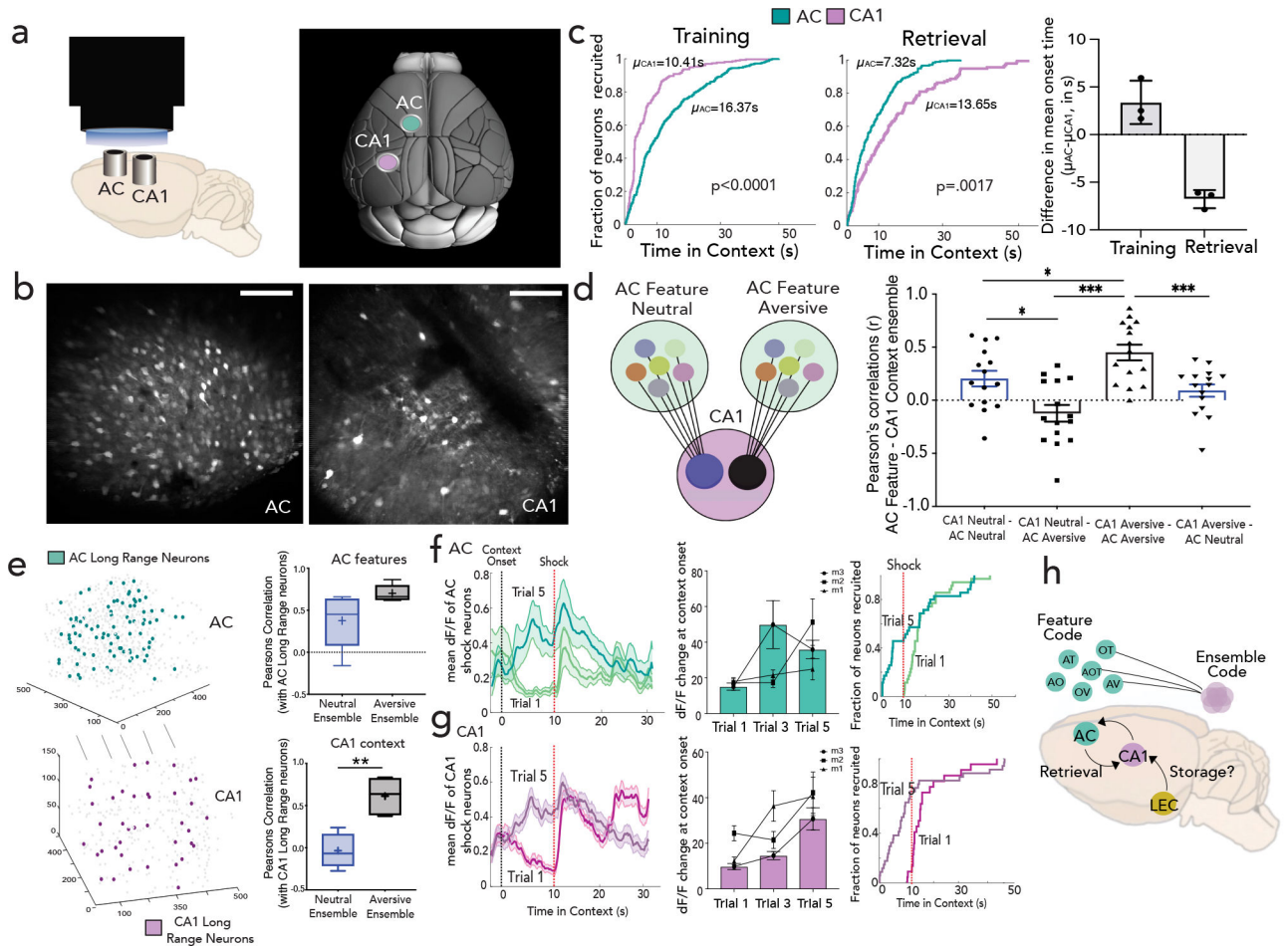


Figure 4 | Dynamic interactions between feature representations in AC and contextual representations in CA1 facilitate memory recall

a-b, Side (left) and top view (right, adapted from Allen Institute Repository) of GRIN lens implantation to access CA1 and AC simultaneously, with Z-projection images (mean over time) from one mouse. Scale: 160 μ m. **c**, Proportion of context selective neurons responding (CDF, cumulative distribution function) to context onset during training as a function of latency (Kolmogorov–Smirnov test, $p < 0.0001$) for a representative mouse (purple-CA1; green-AC). Middle: Same as in (c) but for retrieval, with feature ensembles in AC and context ensembles in CA1 ($p = 0.001$); difference of mean onset time of AC and CA1 during training and retrieval (right) determined by fitting onset curve to an exponential function ($n = 3$ mice, Data are mean \pm s.d.) **d**, Interaction between AC feature and CA1 context ensembles, with dF/F correlation between feature ensembles in AC (all feature types grouped) with context ensembles in CA1 across neutral (blue) and aversive (black) feature trials. (mean \pm s.e.m; $n = 3$ mice, 5 retrieval sessions, 3 features each, Kruskal-Wallis One-Way ANOVA with Dunn’s post-hoc) **e**, Highly connected AC-CA1 neurons (long range partners, left) and their correlations with feature neurons in AC (all feature types grouped, $p = 0.055$) and context neurons in CA1 are quantified ($n = 3$ mice, 5 sessions, Student’s paired t test $**p = 0.009$) **f**, Average activity (shaded bars as s.e.m) of AC shock responsive neurons on trial 1 and trial 5 (black dotted line: context onset, red dotted line: first shock delivery)

for a representative mouse, with quantification of mean dF/F (during first 10 seconds in context normalized to ITI) across all mice ($n=3$ mice, data are $\text{mean} \pm \text{s.e.m}$) (middle); Right: proportion of shock responsive neurons active as a function of time in context in trials 1 and 5 ($n=3$ mice). **g**, Same as in (f) but for CA1. **h**, schematic of working model where LEC inputs to CA1 may be a dedicated storage circuit, whereas ACC inputs to CA1 may be a dedicated retrieval circuit. Details of statistical analyses in Supplementary Table

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript