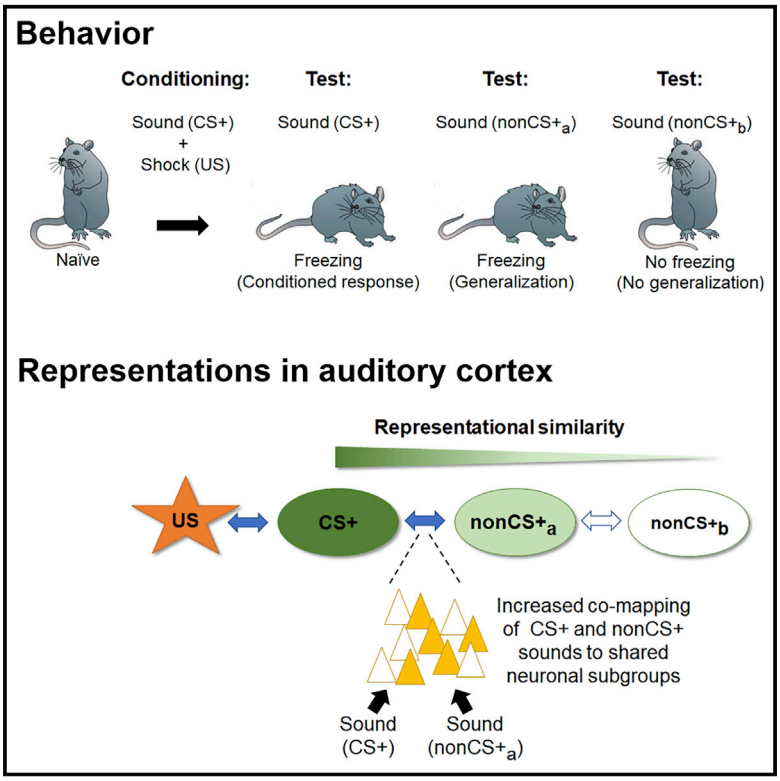
Cell Reports

Article

Learning-induced biases in the ongoing dynamics of sensory representations predict stimulus generalization

Graphical abstract



Highlights

- Chronic *in vivo* two-photon calcium imaging of >36k neurons in mouse auditory cortex
- Drift of auditory representations despite stable environmental conditions
- Increased co-activation of neurons by CS+ and nonCS+ sounds after conditioning
- Degree of co-activation is predictive of behavioral generalization

Authors

Dominik F. Aschauer, Jens-Bastian Eppler, Luke Ewig, Anna R. Chambers, Christoph Pokorny, Matthias Kaschube, Simon Rumpel

Correspondence

sirumpel@uni-mainz.de (S.R.), kaschube@fias.uni-frankfurt.de (M.K.)

In brief

Using longitudinal calcium imaging of sound-evoked activity patterns in the mouse auditory cortex, Aschauer et al. observe increased formation of associations between sound representations following auditory-cued fear conditioning. The degree of newly formed associations predicted behavioral generalization between conditioned and non-conditioned sounds.

Aschauer et al., 2022, Cell Reports *38*, 110340
February 8, 2022
https://doi.org/10.1016/j.celrep.2022.110340

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Learning-induced biases in the ongoing dynamics of sensory representations predict stimulus generalization

Dominik F. Aschauer,[1,6] Jens-Bastian Eppler,[2,6] Luke Ewig,[2] Anna R. Chambers,[1,4] Christoph Pokorny,[3,5] Matthias Kaschube,[2,7,*] and Simon Rumpel[1,7,8,*]

[1]Institute of Physiology, Focus Program Translational Neurosciences, University Medical Center, Johannes Gutenberg University-Mainz, Mainz, Germany
[2]Frankfurt Institute for Advanced Studies and Institute for Computer Science, Goethe University Frankfurt, Frankfurt, Germany
[3]Institute of Theoretical Computer Science, Graz University of Technology, Graz, Austria
[4]Present address: Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway
[5]Present address: Blue Brain Project Connectomics, École Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland
[6]These authors contributed equally
[7]Senior author
[8]Lead contact
*Correspondence: sirumpel@uni-mainz.de (S.R.), kaschube@fias.uni-frankfurt.de (M.K.)
https://doi.org/10.1016/j.celrep.2022.110340

SUMMARY

Sensory stimuli have long been thought to be represented in the brain as activity patterns of specific neuronal assemblies. However, we still know relatively little about the long-term dynamics of sensory representations. Using chronic *in vivo* calcium imaging in the mouse auditory cortex, we find that sensory representations undergo continuous recombination, even under behaviorally stable conditions. Auditory cued fear conditioning introduces a bias into these ongoing dynamics, resulting in a long-lasting increase in the number of stimuli activating the same subset of neurons. This plasticity is specific for stimuli sharing representational similarity to the conditioned sound prior to conditioning and predicts behaviorally observed stimulus generalization. Our findings demonstrate that learning-induced plasticity leading to a representational linkage between the conditioned stimulus and non-conditioned stimuli weaves into ongoing dynamics of the brain rather than acting on an otherwise static substrate.

INTRODUCTION

It is believed that sensory-evoked activity patterns at the level of sensory cortices can serve as a neural correlate of a percept. In light of the common day experience that the same sensory stimulus evokes the same percept from day to day, the most parsimonious assumption is that the underlying sensory representations are also stable. However, over the last years, there is growing evidence that this assumption may not always be justified and that neuronal tuning to sensory stimuli undergoes ongoing remodeling even under behaviorally and environmentally stable conditions: long-term remodeling of functional properties of neurons has been reported previously in the mouse hippocampus and barrel, visual, motor, and posterior parietal cortex (Clopath et al., 2017; Deitch et al., 2021; Driscoll et al., 2017; Hainmueller and Bartos, 2018; Huber et al., 2012; Mankin et al., 2012; Margolis et al., 2012; Rokni et al., 2007; Rule et al., 2019; Schoonover et al., 2021; Ziv et al., 2013). However, it is still unclear how functional changes at the level of individual neurons affect the properties of sensory representations at the population level and how such ongoing changes relate to those expected to occur during learning. This gap in knowledge is not only due to the scarcity of data up to now, but is also a consequence of the difficulty to establish appropriate frameworks to describe complex population dynamics in a biologically interpretable manner.

The mouse auditory cortex offers an interesting model to study the longevity of sensory representations. It has been previously shown that sounds lead to the non-linear co-activation of neuronal subgroups at the local microcircuit scale (Atencio and Schreiner, 2013; Bathellier et al., 2012; See et al., 2018). A given, specific local neuronal subgroup can be activated by a subset of different sounds and therefore gives rise to near identical response patterns. We called the stereotypical local response pattern that is associated with this subset of sounds in a population of imaged neurons a *response mode* (Bathellier et al., 2012). The combination of different sounds that are mapped to a specific response mode varies across the auditory cortex. Thus, a distinct representation for a specific sound is formed at the global scale by the combinatorial pattern of multiple local response modes. Furthermore, this global description of neuronal population activity based on local response modes is

sufficient to predict spontaneous categorization behavior in mice trained to discriminate pairs of sound stimuli (Bathellier et al., 2012). The discrete nature of auditory representations at the population level provides a highly sensitive and robust readout for changes occurring over the time course of days and enables a systematic assessment of their dynamics.

Here, we used chronic two-photon calcium imaging in the mouse auditory cortex over several days to monitor sound-evoked activity patterns forming sensory representations. We find that, even in behaviorally habituated mice, auditory representations display significant plasticity on the single-cell and population levels. Exploiting the discrete nature of response modes, we observed near stationary dynamics involving substantial remapping of stimuli to response modes as well as the creation and elimination of individual response modes, while leaving the statistics of neural activity patterns observed at a single time point largely unaffected. Applying the same analysis to data from mice undergoing auditory-cued fear conditioning (ACFC), we observed specific biases in the mapping of stimuli to response modes resulting in an increased co-activation of shared subgroups of neurons activated by the conditioned stimulus and specific non-conditioned stimuli that are characterized by a high representational similarity to the conditioned stimulus. These dynamics of enhanced co-mapping suggest the formation of new sensory associations to the conditioned stimulus and were predictive of the level of stimulus generalization we observed behaviorally.

RESULTS

Chronic large-scale calcium imaging of neurons in the mouse auditory cortex

To assess the long-term dynamics of auditory representations, we transduced cells in the mouse auditory cortex with a co-injection of two rAAV8 vectors to drive stable expression of two fluorescent proteins under the control of the synapsin promoter: the genetically encoded calcium indicator GCaMP6m (Chen et al., 2013), to chronically record neural activity, and the fusion protein H2B:mCherry, as a structural marker to distinctively label the nuclei of transduced neurons (Figures 1A–1C) (Nathanson et al., 2009).

We used intrinsic signal imaging in response to a set of pure-tone stimuli of varying frequency to guide subsequent two-photon imaging. For calcium imaging experiments in awake, head-fixed, passively listening mice, we used a stimulus set of brief (50–70 ms) sounds containing 19 sinusoidal pure tones and 15 complex sounds characterized by temporally modulated power in multiple frequency bands delivered free field using a calibrated speaker at 74 dB sound pressure level (Figure S1A).

Mice were habituated to head fixation and pre-exposed to the set of sound stimuli for at least 5 days to ensure that adaption to novel sensory responses was largely completed (Kato et al., 2015) and that data acquisition occurred under behaviorally and environmentally familiar and constant conditions. The red nuclear marker enabled high-fidelity re-identification and registration of individual local populations that were re-imaged for four time points at a 2-day interval (Figures S2A–S2D). We imaged neuronal activity in response to the 34 sound stimuli (20–30 repetitions each, presented in a pseudorandom order) in a total number of 21.506 neurons in 97 different fields of view (~100–300 neurons per FOV) in the cortical layer 2/3 of 12 mice (Figures 1D and S2E–S2G). When assessing trial-averaged calcium responses to pure tones and complex sounds over this period, we observed that many neurons that were responsive within an FOV showed essentially stable responses to the sound stimuli over a 2-day interval (30.75% ± 0.01% of responses). Others, however, showed substantial re-tuning involving the gain (36.85% ± 0.02%) or the loss (32.41% ± 0.01%) of responses with substantial signal amplitudes (Figures 1E and S1B).

Dynamic long-term remodeling of sound responses in individual neurons and populations of neurons

In light of the re-tuning of sound responsiveness in a substantial fraction of neurons, we next asked how these changes would affect the ability of the auditory cortex to form a stable representation of the auditory world. When pooling the data obtained from all mice and FOVs, we identified on any given day a comparable number of neurons showing significant responses for any of the 34 stimuli used in this study (Figures 2 and S3). However, when considering only those neurons that displayed significant sound-evoked responses on the first imaging day, and following them across days, we observed a progressive blurring of the response profile with time, reflecting the fact that some neurons changed their preferred stimulus or became unresponsive. This process was largely mirrored when considering only those cells with sound-evoked responses on the last imaging day, highlighting neurons that gained responsiveness to sounds during the course of the experiment. We quantified the instability of responses across days by computing the averaged, normalized response amplitude for the preferred stimulus. If computed for the stimulus that was driving the neurons' maximal response on that given day, this quantity is 1 and stable, by definition. However, if this analysis is performed on each day for the preferred stimulus from the first day, we found a substantial and continuous loss of average response amplitudes. Again, a symmetric observation was made when normalizing to the preferred stimulus amplitude of the last day. The observation that, despite substantial changes in sound tuning of individual cells, a similar number of cells is sound responsive on a given day, indicates that representations in the auditory cortex are maintained in a dynamic equilibrium at a global level.

Next, we asked how these changes, observed on the level of individual cells, become manifest on the level of population activity. We observed that population responses were often stable across several days. However, consistent with our single-cell analysis above, sometimes the set of neurons responding to a given stimulus changed from one imaging day to the next (Figures 3A–3C, S4A, and S4B). To assess whether these changes affect the ability of neural populations to stably distinguish between auditory stimuli, we trained a linear classifier to discriminate single-trial activity patterns elicited by different sound stimuli in a given FOV (see STAR Methods). When training and testing with activity patterns recorded at the same day, we observed similar performances across different imaging days (Figure 3D). However, the impact of the changes in the

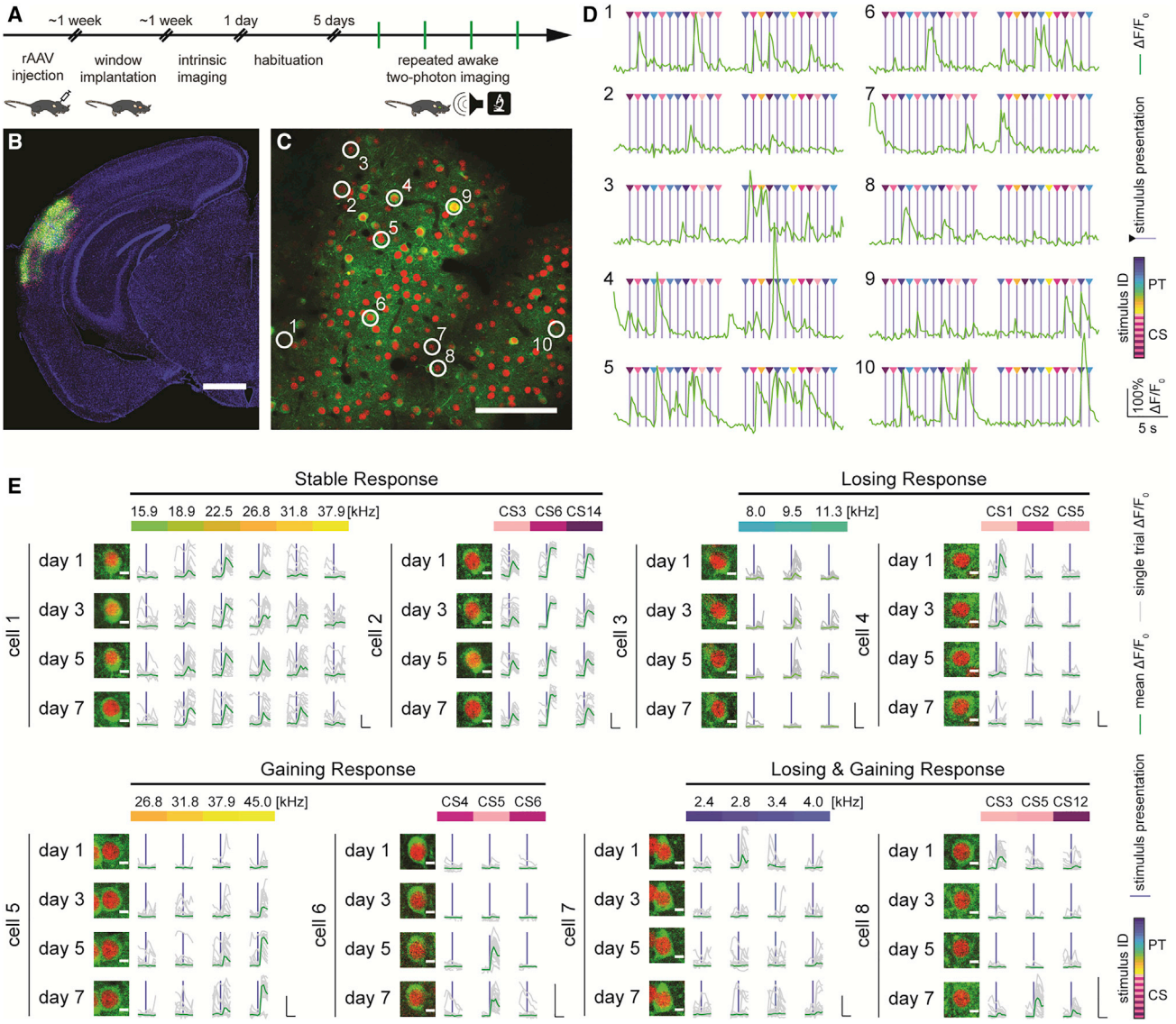


Figure 1. Chronic imaging of sound-evoked activity in auditory cortex of awake mice

(A) Experimental timeline. Dataset comprises 21,506 cells from 12 mice.

(B) Confocal image of a coronal section of a mouse brain sacrificed 46 days after stereotactic injection of a mixture of two rAAVs leading to the expression of GCaMP6m (green) and H2B::mCherry (red) in the auditory cortex. Counterstain is DAPI (blue). Scale bar, 1 mm.

(C) *In vivo* two-photon image of a local population of neurons in layer 2/3 of the auditory cortex showing expression of GCaMP6m (green) and H2B:mCherry (red). Scale bar, 100 μm. White circles and digits represent the location of the neurons in (D).

(D) Simultaneously recorded activity traces of 10 example neurons from the two-photon field of the view shown in (C) during awake, passive listening (green, $\Delta F/F_0$; blue line, stimulus presentation; colored arrowheads, stimulus identity; PT, pure tones; CS, complex sounds).

(E) Responses to auditory stimuli in single neurons monitored over multiple days (green, mean $\Delta F/F_0$; gray, single trial $\Delta F/F_0$; blue line, stimulus presentation; insets show images of cells on different days; image scale bar, 5 μm; trace scale bar, 1 s, 250% $\Delta F/F_0$). Note that individual cells show both stable and plastic features in tuning.

population response on stimulus discrimination became evident when training the classifier on sound responses from the first day and testing the performance with activity patterns from the following imaging days. The decoding performance decreased monotonically with an increasing interval between training and testing. As above, we made a symmetrical observation when training the classifier on the data from the last imaging day and testing with sound responses from previous imaging days. Consistent observations were made using an alternative, multi-class decoding approach (Figure S4C). Different subsets of stimuli evoked activity patterns in the various FOVs within the same mouse (Bathellier et al., 2012). Therefore, combining the information from various FOVs led to a substantial increase in decoding compared with single FOVs (Figure S4D). Thus,

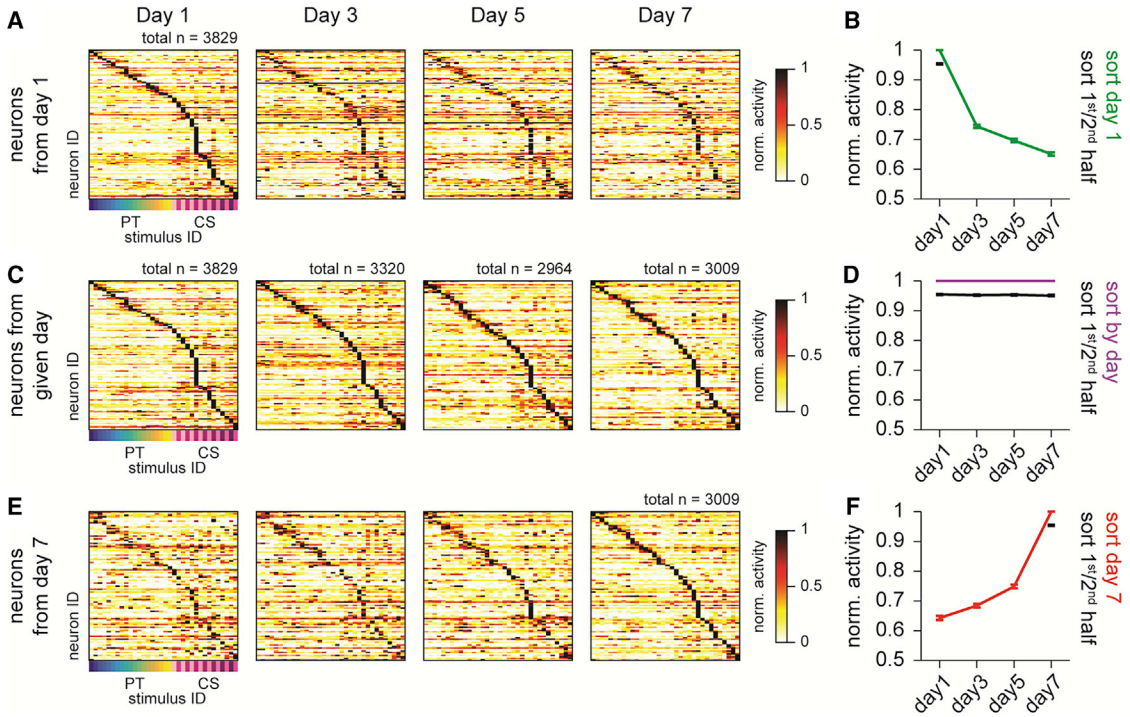


Figure 2. Balanced drifts in tuning at the single-cell level

(A) Normalized response profiles of neurons with a significant response to at least one stimulus on day 1 sorted by stimulus with highest response amplitude. Sorting from day 1 is applied to the subsequent days. N is total number of significantly sound-responsive cells on each day. For illustrative purposes, only every 30th cell is shown (PT, pure tones; CS, complex sounds).

(B) Average (mean ± SEM) normalized activity to the stimulus with highest response amplitude on day 1 (A) plotted across days. Estimation of best stimulus is robust against sub-sampling of trials (black; Figure S3).

(C) Same as (A), for cells with a significant response on a given day. Sorting is done for each day individually.

(D) Same as (B), for cells with a significant response on a given day.

(E) Same as (A), for cells with a significant response on day 7. Sorting from day 7 is applied to the previous days.

(F) Same as (B), for cells with a significant response on day 7.

extending the analysis to the level of neuronal populations, we found that the ability to decode sounds from the auditory cortex was largely robust against the ongoing remodeling of sound-evoked response patterns.

A discrete set of response modes forms a population-level representation of the auditory world

Having observed a considerable degree of ongoing remodeling, not only of tuning properties in individual neurons but also of neural representations on the population level, we next sought to capture the essence of these changes more efficiently. To this end, we exploited the fact that, on a microcircuit scale, sound-evoked activity patterns fall into a near-discrete set of response modes (Bathellier et al., 2012).

This phenomenon is illustrated for an example FOV in Figure 4A, where we probed the population response with a set of pure tones with gradually changing frequency. Instead of a unique and distinct population response for each of the pure tones, we observed an abrupt, highly non-linear transition from one response pattern to another. To further assess the similarity of the response patterns, we constructed a similarity matrix by calculating pairwise correlations of single-trial response vectors. The entries along the diagonal were calculated as the average Pearson correlation of all pairwise combinations of response vectors elicited by the same sound, thus reflecting the trial to trial reliability of the response pattern elicited by a given sound. Correlation was low for sounds that did not evoke any activity in the particular FOV. Off-diagonal entries were calculated as the average correlation of all pairwise combinations of single-trial response vectors elicited by a pair of two different sounds (see STAR Methods). The two prominent clusters highlighted that, in this example, the stimuli from 2 to 4 kHz evoked highly similar activity patterns, allowing to describe their collective stereotypical response pattern by a single response mode (Figure 4B). Likewise, the activity patterns elicited by the stimuli between 4 and 10 kHz showed a high degree of similarity and therefore constituted a second response mode. The stimuli 11.3 to 15.9 kHz evoked no response in this FOV and were assigned to the so-called 0-mode.

Applying the correlation analysis together with subsequent clustering (Figure S5) we observed, in most FOVs in our dataset, that the activity patterns elicited by subsets of stimuli could be well described by a shared, stereotypical response mode. The composition of the subsets of sound stimuli mapped to a shared

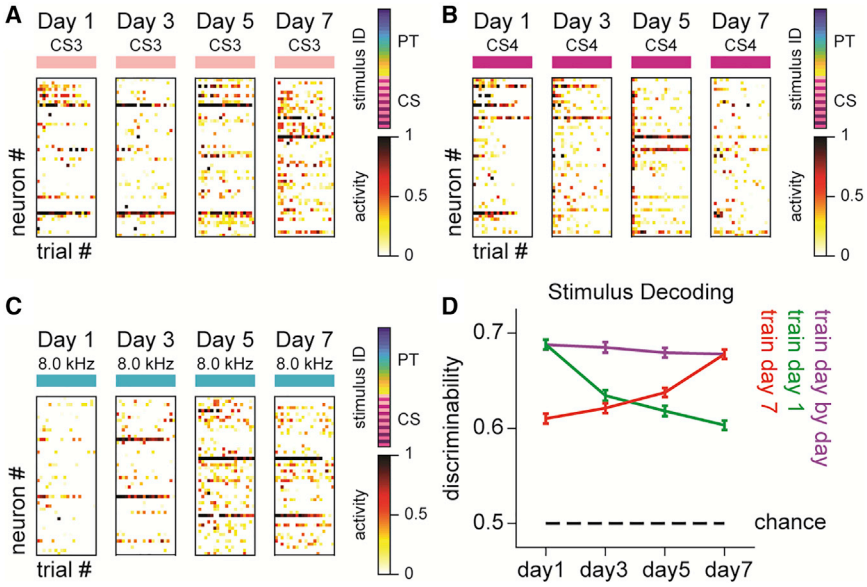


Figure 3. Population responses to auditory stimuli are dynamic under basal conditions

(A–C) Single-trial population response vectors acquired from a given FOV. Examples are shown for three example stimuli and three different FOVs over the time course of 7 days. For illustrative purposes, only the 50 most active cells are shown in random order and trials are sorted by descending mean activity (PT, pure tones; CS, complex sounds).

(D) Linear pairwise discriminability calculated by logistic regression averaged across all possible sound pairs and FOVs (mean ± SEM) plotted across days. The classifier was trained with data from either first (green), last (red), or given (purple) imaging day. Dashed line indicates chance level.

response mode varied across FOVs and often comprised a mix of both pure tones and complex sounds. In a given FOV we typically observed between three and nine different response modes, on average associated with 2.40 ± 0.21 (mean ± SEM) sound stimuli and comprising in the majority of FOVs less than 10 significantly active neurons (Figure S6).

In summary, our analysis of the data based on individual imaging days corroborates previous reports suggesting a functional layout of the superficial layers of mouse auditory cortex with scattered and partially overlapping neuronal subgroups that are driven in a non-linear manner by different groups of sound stimuli (Bathellier et al., 2012; See et al., 2018).

Ongoing recombination of sensory representations during basal conditions

Having observed that the structure of auditory representations in an FOV can be well approximated by a small set of response modes, we exploited this highly reduced, non-linear description to capture the main aspects of the ongoing representational changes during basal conditions. We constructed similarity matrices from the response patterns observed on all imaging days within an FOV and applied hierarchical clustering to statistically identify individual response modes (see STAR Methods and Figure S5). The analysis revealed that, over the time course of several days, the response mode structure forming a sound representation undergoes substantial remodeling. This dynamics becomes obvious when considering the clustered similarity matrices of the response patterns for the four imaging days sorted on each day individually or even more dramatically when keeping the sorting of the first or last day fixed for the remaining days (Figures 5A–5C and S7). Here, the response modes that, after sorting, became apparent as few, well-defined clusters in the matrices, gradually eroded over the time course of several days.

To systematically analyze these ongoing changes, we assessed for each of the 97 FOVs and each of the 34 stimuli if they were mapped to any response mode. We found that, on a given day, about a third of the stimuli was mapped to a response mode, but over the time course of a week, about half of the stimuli elicited at least once a population response (Figure 5D). In principle, these changes could be explained by unstable mapping of individual stimuli to a set of stable response modes. However, when quantifying the number of response modes observed in the 97 FOVs, we also observed that response modes themselves underwent ongoing remodeling, i.e., new response modes were identified at a later time point, whereas others were no longer detected (Figure 5E). Of all stimuli being mapped to a specific response mode on a given day, only half remained in that mode on the following imaging day, whereas almost 20% moved to a different response mode, and more than 30% to the 0-mode, i.e., no longer elicited a population response (Figure 5F). This dynamics was largely balanced, as almost 30% of stimuli mapped to a response mode on a given day did not elicit a population response on the previous imaging day. In summary, we found changes in the mapping of sounds to response modes, as well as changes in the response modes themselves, while the average number of stimuli being mapped to a response mode and the total number of response modes remained fairly stable across imaging days.

ACFC increases population responses for stimuli with representational similarity to the conditioned sound

We next wondered to what extent behaviorally relevant experiences that trigger the formation of a memory to a sound would impact on the long-term dynamics of sensory representations in mouse auditory cortex (Aschauer and Rumpel, 2018). To this end, we acquired a second dataset (10 mice; 74 FOVs; 16,882 neurons) with four imaging time points at a 2-day interval using the same set of sound stimuli as before (Figure 6A). On the day between imaging sessions two and three, mice underwent an ACFC paradigm in which they learned to associate the presentation of a sound with the subsequent application of a mild foot shock. A complex sound (CS6) from the stimulus set was chosen as conditioned stimulus (CS+) for all conditioning experiments. It has previously been shown that ACFC and variants thereof

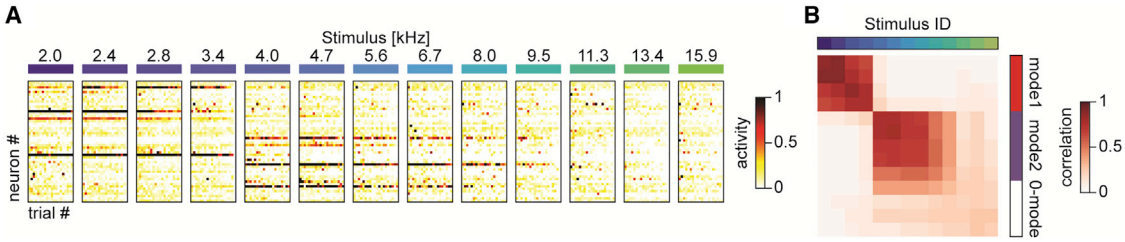


Figure 4. Abrupt transitions in population responses elicited by gradually changing stimuli delineate discrete response modes

(A) Example population activity from an FOV showing a non-linear response mode transition to interpolation of pure-tone frequencies from 2 to 15.9 kHz. Top: stimulus identity; bottom: single trial population response vectors. For illustrative purposes, only the 50 most active cells are shown in random order and trials are sorted by descending mean activity.

(B) Similarity matrix (mean Pearson correlation) from the example FOV shown in (A). Top: stimulus identity; right: response mode identity; 0-mode: no significant population response (see STAR Methods).

induce nucleus-wide reorganization of chromatin and specific changes in gene expression (Cho et al., 2017; Peter et al., 2012, 2021), affect the dynamics of synaptic connections (Lai et al., 2018; Moczulska et al., 2013; Yang et al., 2016), and induce changes in the tuning of neurons in the mouse auditory cortex (Dalmay et al., 2019; Gillet et al., 2017; Quirk et al., 1997; Weinberger, 2004).

To test for the successful formation of a memory at the behavioral level, we exposed mice again to the sound cue used for conditioning in a neutral context after the last imaging session and scored freezing behavior as a readout of fear-related memory recall. As expected, we observed low freezing levels during silence and significantly increased freezing during presentation of the conditioned sound. Furthermore, we also observed high freezing levels during the presentation of a second sound stimulus that was not presented during the conditioning session (nonCS+ (a)) (Figure 6B; n = 10 mice; mean ± SEM: first 30 s: 5.00 ± 2.51, CS+: 28.49 ± 3.73, nonCS+: 34.26 ± 3.45; df: 2; $F = 22.47$; *$p < 0.0001$ for first 30 s and CS+, *$p < 0.0001$ for first 30 s and nonCS+). The observation that non-conditioned stimuli sharing perceptual similarity with the CS+ can also evoke conditioned responses is commonly described as stimulus generalization (Pavlov and Anrep, 1927). A high level of generalization is typically observed following classical fear conditioning unless specific differential conditioning paradigms are utilized (Letzkus et al., 2011).

We compared the fraction of sound-responsive cells in both datasets and found only a transient increase in the first imaging session after conditioning (Figure 6C; control: n = 97 FOVs; mean ± SEM; 1 ± 0, 0.989 ± 0.061, 0.916 ± 0.057, 0.933 ± 0.046, ACFC: n = 74 FOVs; mean ± SEM; 1 ± 0, 1.057 ± 0.047, 1.110 ± 0.055, 0.971 ± 0.039; *$p < 0.0001$). The efficacy to decode sounds from population response vectors recorded within individual FOVs; however, showed a long-lasting increase following conditioning in comparison with the control group (Figure 6D; control: n = 97 FOVs; mean ± SEM; 0.6878 ± 0.0052, 0.6849 ± 0.0056, 0.6794 ± 0.0051, 0.6779 ± 0.0049, ACFC: n = 74 FOVs; mean ± SEM; 0.6866 ± 0.0059, 0.6824 ± 0.0056, 0.6866 ± 0.0055, 0.6874 ± 0.0061; *$p < 0.05$). In parallel to this, the average number of stimuli that evoked a population response in a given FOV, i.e., were mapped to a response mode, was significantly increased following conditioning compared with control (Figure 6E; control: n = 97 FOVs; mean ± SEM 11.27 ± 0.68, 10.95 ± 0.70, 9.90 ± 0.59, 9.12 ± 0.57, ACFC: n = 74 FOVs; mean ± SEM 11.78 ± 0.75, 11.01 ± 0.70, 11.01 ± 0.69, 11.22 ± 0.73; day 5: *$p = 0.0095$, day 7: *$p < 0.0001$).

We next wondered to what extent the response increase observed for a given stimulus is predicted by its similarity to the CS+ prior to conditioning. As our knowledge of the relevant perceptual space in mice is very limited at present, we used representational similarity in the auditory cortex as a proxy for stimulus similarity. We followed an approach to use the correlation of stimulus-evoked activity patterns to gain a pairwise measure of the representational similarity between the 34 sounds including the CS+ (Kriegeskorte et al., 2008). We constructed correlation matrices by calculating the Pearson correlation between the pairwise response vectors for all sound stimuli recorded on the first imaging day in each FOV of both datasets and then averaged them across all FOVs (Figure 6F). This yielded a robust measure of representational similarity, as we obtained highly consistent estimates for both datasets analyzed separately (Figure S8). We next sorted the stimuli according to their similarity to the CS+ and calculated for each dataset the fraction of FOVs in which a particular nonCS+ stimulus elicited a population response. During baseline (day 1), there was no systematic difference between the ACFC and control datasets (Figure 6G; gray dots, Spearman rank correlation $\rho = 0.32$, $p = 0.07$). In contrast, 3 days after conditioning (day 7) we found that a subset of stimuli were more likely to elicit a population response in the conditioned group. This observation was specific for stimuli with a high representational similarity to the CS+, as other stimuli with low similarity to the CS+ showed only minor changes (Figure 6G; black dots, Spearman rank correlation $\rho = 0.66$, $p < 0.0001$). Notably, responses to the CS+ itself seemed to be barely affected by conditioning.

Fear conditioning drives the formation of associations between sensory representations, predicting behavioral generalization

Stimulus generalization is believed to involve the linkage or association between the conditioned CS+ and other nonCS+ stimuli during conditioning (Dunsmoor and Paz, 2015; Pavlov and Anrep, 1927). At the microcircuit level, the activation of a shared subgroup of neurons by different stimuli has been interpreted

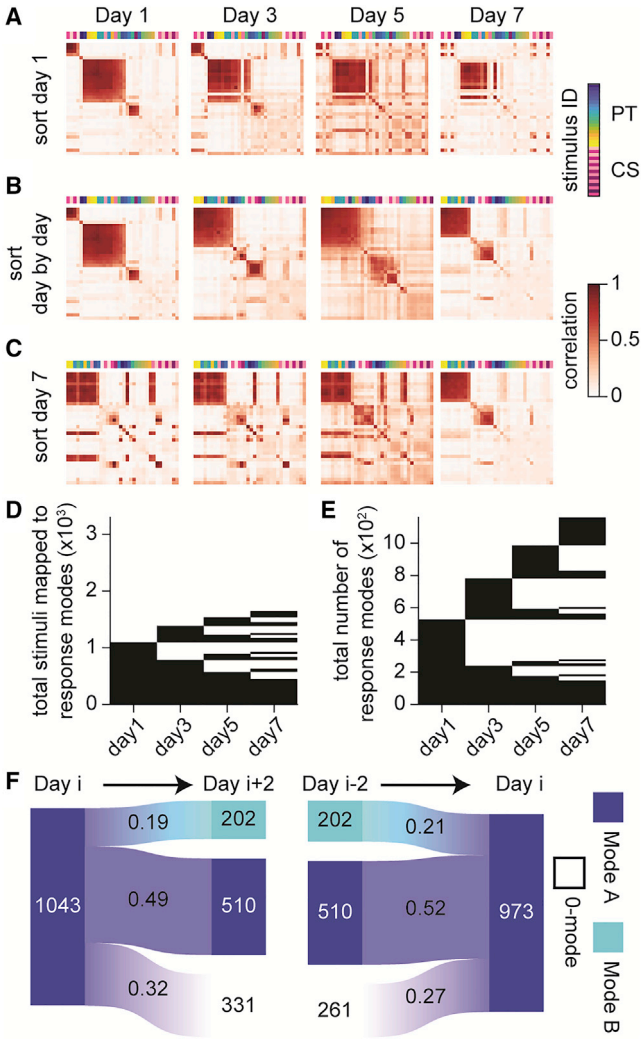


Figure 5. Recombination of response modes over the course of days

(A) Single-day similarity matrices of sound-evoked responses from an example FOV sorted by hierarchical clustering. The clustering procedure sorts the sound stimuli according to the similarity of their corresponding response patterns observed in the FOV. Sorting from day 1 is applied to the subsequent days. Top: stimulus identity (PT, pure tones; CS, complex sounds).

(B) Same as (A) but sorted on each day individually.

(C) Same as (A) but sorting from day 7 is applied to the previous days.

(D) Life-time plot depicting the stability of the mapping of stimuli to response modes. The mapping was assessed for the 34 stimuli in all 97 FOVs resulting in a maximally possible number of 3.298 mappings, displayed along the y axis. A thin horizontal black line indicates a mapping (significant response) present on a given day. Mappings are sorted based on day 1, from present to non-present mappings. Data from 12 mice.

(E) Life-time plot of the total number of response modes that were identified in a given FOV. A thin horizontal black line indicates presence of a response mode on a given day, data pooled over all 97 FOVs and 12 mice.

(F) Flow chart of response mode dynamics. Left: development of population responses present on day i (defined as mode A) 2 days into the future (mode A, mode B, or 0-mode). Right: development of population responses present on day i (defined as mode A) from 2 days previously (mode A, mode B or 0-mode). Numbers are counts and fractions averaged across transitions.

as the formation of an association among sensory representations (Grewe et al., 2017). We therefore leveraged our framework to describe population dynamics using response modes to test if ACFC leads to changes consistent with an increased formation of associations between CS+ and nonCS+ representations. In this framework the formation of associations would be reflected by an increased number of stimuli being mapped to the same response mode. Such a change would be consistent with our observation that fear conditioning increases the number of sounds driving a population response in a given FOV as described above (Figures 6E and 6G). Alternatively, however, an increased likelihood of population responses could also be explained by an increased formation of new response modes, without involving increased co-mapping of stimuli to the same subgroup of neurons.

Our analysis of the response mode dynamics revealed that learning induces a bias in the ongoing recombination of sensory representations compared with basal conditions. Specifically, we found that the average number of modes per FOV was indistinguishable between the baseline and fear conditioning cohorts during all imaging days (Figure 7A; control: n = 97 FOVs; mean ± SEM; 5.42 ± 0.32, 5.10 ± 0.34, 4.93 ± 0.30, 4.69 ± 0.29, ACFC: n = 74 FOVs; mean ± SEM; 5.51 ± 0.37, 5.49 ± 0.35, 4.97 ± 0.32, 5.01 ± 0.31), whereas the average number of stimuli per response mode was significantly larger following conditioning (Figure 7B; control: n = 97 FOVs; mean ± SEM; 2.39 ± 0.21, 2.27 ± 0.13, 2.19 ± 0.13, 1.99 ± 0.11, ACFC: n = 74 FOVs; mean ± SEM; 2.36 ± 0.12, 2.11 ± 0.13, 2.48 ± 0.17, 2.53 ± 0.20; day 5: *p = 0.0021; day 7: *p < 0.0001). Both datasets in our study were dominated by a substantial degree of ongoing representational dynamics (Figure S9). Under basal conditions, we observed that stimuli disappeared from a given response mode, by a rate that was almost balanced by new stimuli being added to it (Figure 5F). Intriguingly, we found that, after fear conditioning, both of these rates were shifted: the rate of stimuli being added to a mode per day increased, relative to the baseline dynamics (Figure 7C; control: n = 97 FOVs; mean ± SEM; 1 ± 0.056, 0.902 ± 0.054, 0.829 ± 0.052, ACFC: n = 74 FOVs; mean ± SEM; 1 ± 0.065, 1.032 ± 0.066, 1.027 ± 0.065; transition 3→5: *p < 0.0001; transition 5→7: *p < 0.0001), while the rate of stimuli disappearing (and entering the 0-mode) decreased (Figure 7D; control: n = 97 FOVs; mean ± SEM; 1 ± 0.054, 1.186 ± 0.058, 1 ± 0.054, ACFC: n = 74 FOVs; mean ± SEM; 1 ± 0.057, 0.819 ± 0.052, 0.761 ± 0.05; transition 3→5: *p < 0.0001; transition 5→7: *p < 0.0001). Both processes effectively increased the average number of stimuli that are mapped to a response mode. The former suggests an increase in the formation of new associations among stimuli, while the latter suggests the stabilization of existing associations. Note that such an increase in the average number of stimuli per response mode is compatible with the overall improvements of stimulus decoding that we observed following fear conditioning, as we illustrated by means of a simple, mathematical model (Figure S10).

We next assessed if increased co-mapping of stimuli following fear conditioning is specific to nonCS+ sound stimuli similar to the CS+. In the basal control group, as expected, sounds with a higher representational similarity before conditioning also had a higher likelihood to be co-mapped to the same response

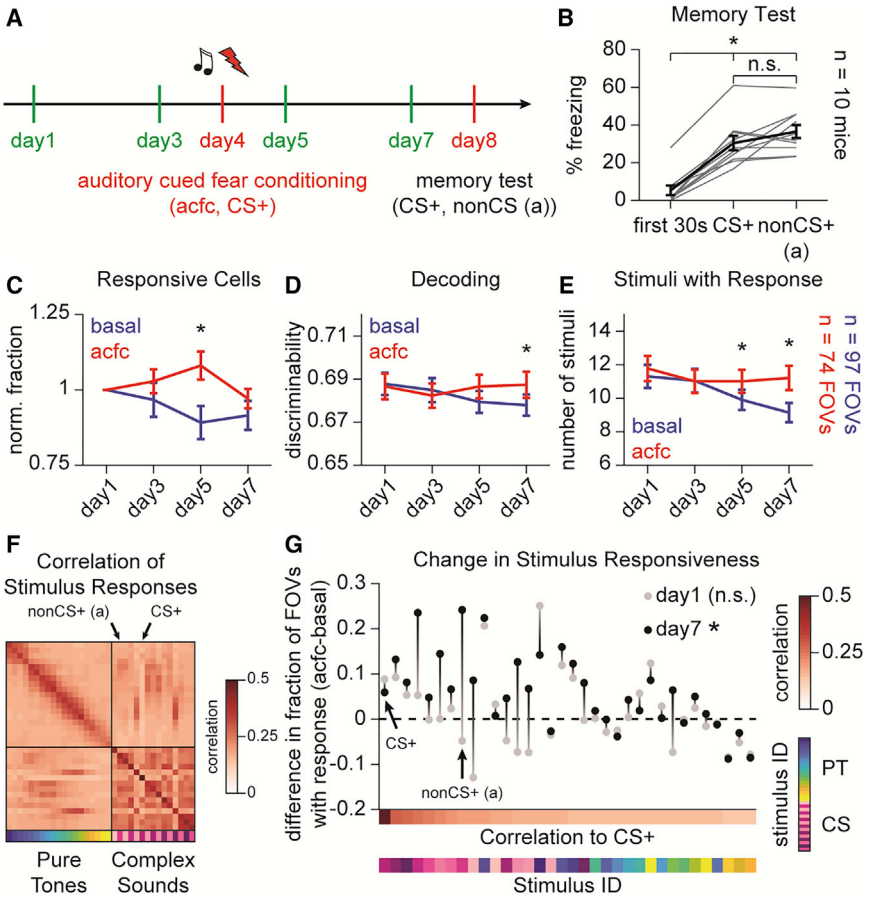


Figure 6. Auditory-cued fear conditioning increases responsiveness for stimuli similar to the CS+

(A) Experimental timeline for cohort of mice undergoing chronic calcium imaging and auditory-cued fear conditioning (ACFC). During conditioning a complex sound stimulus (CS6) was used as conditioned stimulus. In the memory test session mice were exposed to the CS+ as well as another non-conditioned sound (nonCS+). The imaging dataset comprises 16,882 cells from 10 mice.

(B) Increase in freezing behavior for conditioned stimulus (CS+) and high level of generalization to nonCS+ (a) in the memory test 4 days after fear conditioning. Gray lines depict behavior of individual animals and black line is mean ± SEM of all animals (*one-way ANOVA with correction for multiple comparisons, $p < 0.0001$).

(C) Normalized fraction of sound-responsive cells over the imaging days averaged over FOVs (mean ± SEM, *bootstrap test, $p < 0.0001$).

(D) Decoding performance (pairwise logistic regression, training and test data from same day) in baseline and ACFC dataset (mean ± SEM, *bootstrap test, $p < 0.05$).

(E) Average number of stimuli eliciting a population response in a given FOV (mean ± SEM, *bootstrap test, day 5: $p < 0.01$, day 7: $p < 0.0001$).

(F) Correlation matrix of sound response vectors for all stimuli averaged across all FOVs from both datasets on day 1. Arrows mark columns representing the stimulus used for fear conditioning (CS+) and the non-conditioned stimulus presented during the memory test session (nonCS+ (a)).

(G) Difference between basal and fear conditioning groups in the fraction of FOVs in which a population response to a given stimulus was observed. Gray dots mark values for baseline (day 1) and black dots after conditioning (day 7). Stimuli are sorted on the x axis by descending similarity to the stimulus used for fear conditioning. Correlation color bar represents the correlation of population response vectors to the CS+ (see CS+ column) from (F). Arrows mark the stimulus used for fear conditioning (CS+) and the non-conditioned stimulus (nonCS+ (a)) presented during the memory test session. Note the strong increase in responsiveness after fear conditioning for stimuli with a high correlation to the CS+. *Spearman's rank correlation, $\rho = 0.66$, $p < 0.0001$.

mode, as both measures of population activity are related to one another. Importantly, when comparing the basal control group to the dataset obtained from conditioned mice, we found that the likelihood of co-mapping to the CS+ was more pronounced specifically for the non-conditioned sounds with a high representational similarity to the CS+ (Spearman rank correlation: $\rho = 0.73$, $p < 0.0001$; Figures 7E, 7F, and S11). This suggests that fear conditioning led to an increased formation of associations between the CS+ and those nonCS+ sounds that a priori showed a high level of representational similarity.

To test the correlation between the changes at the level of cortical representations and behavioral stimulus generalization, we performed another fear-conditioning experiment. Here, mice were conditioned to the same complex sound stimulus CS6 as CS+ and we used three nonCS+ sounds to probe for stimulus generalization in a test session 4 days later. We selected two sounds from our set of 34 stimuli, one that showed a high increase of co-mapping to the CS+ (nonCS+ (b)) and one that showed no change (nonCS+ (c)) (Figure 7F). In addition, we included the non-conditioned stimulus used in our previous experiment (nonCS+ (a)) to replicate our finding. Importantly, all stimuli used in this experiment did not induce freezing in naïve mice (Figure S12). We found that, after conditioning, freezing levels induced by nonCS+ (a) and nonCS+ (b) were significantly higher than those observed in response to nonCS+ (c) (Figure 7G; n = 16 mice; mean ± SEM: first 30s: 22.50 ± 2.28, nonCS+ (a): 63.05 ± 3.41, nonCS+ (b): 68.48 ± 2.38, nonCS+ (c): 41.80 ± 4.10, blank: 37.21 ± 4.37; df: 4; F = 30.87; *$p < 0.0001$ for first 30 s and all other groups, *$p < 0.0001$ for nonCS+ (a, b) and nonCS+ (c) and blank). Together, our findings indicate that representational similarity is predictive of the degree of conditioning-induced mapping of nonCS+ stimuli onto the same response mode as the CS+ and that this co-mapping, in turn, can predict the degree of behavioral generalization to a non-conditioned stimulus.

DISCUSSION

We studied the dynamics of auditory representations in the cortex in thousands of neurons over the time course of several days. To

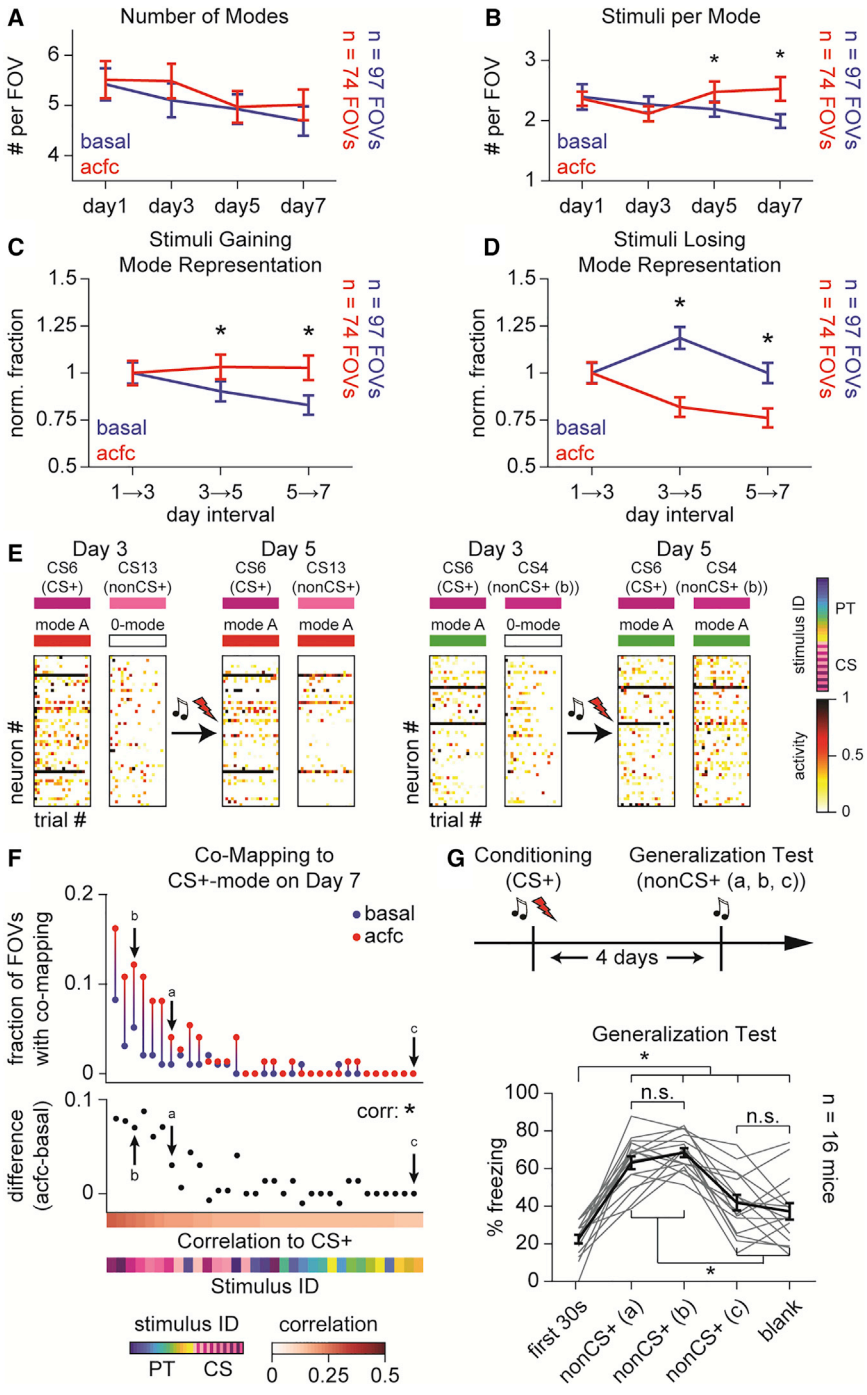


Figure 7. Increased co-mapping of the CS+ and nonCS+ stimuli to shared response modes predicts behavioral generalization

(A) Mean number of response modes averaged over FOVs (mean ± SEM) in both experimental groups.

(B) Mean number of stimuli per response mode averaged over FOVs (mean ± SEM). *Bootstrap test, day 5: $p < 0.01$, day 7: $p < 0.0001$.

(C) Normalized fraction of stimuli gaining a response mode representation averaged over FOVs (mean ± SEM). *Bootstrap test, days 3–5: $p < 0.0001$, day 5–7: $p < 0.0001$.

(D) Same as (B) for stimuli losing a response mode representation. *Bootstrap test, days 3–5: $p < 0.0001$, day 5–7: $p < 0.0001$.

(E) Examples from two FOVs showing the responses of the conditioned stimulus (CS+) and a non-conditioned stimulus (nonCS+). Prior to fear conditioning the nonCS+ did not elicit a significant response (0-mode), whereas after fear conditioning its response became similar to that of the CS+ (mode A). Top: stimulus identity; middle: mode identity; bottom: single trial population response vectors. For illustrative purposes, the 50 most active cells are shown in random order and trials are sorted by descending mean activity (PT, pure tones; CS, complex sounds).

(F) Top: for each stimulus, sorted on the x axis by descending similarity to CS+, the plot shows the fraction of FOVs this stimulus is co-mapped to the response mode of the CS+ on day 7 in experimental groups with (acfc, red) and without (basal, blue) fear conditioning. Correlation color bar represents average correlation of population response vectors to the CS+ (same as in Figure 6G). Arrows mark non-conditioned stimuli presented during the memory test session in (G) (nonCS+ (a, b, c)). Bottom: differences between fractions for acfc and basal. *Spearman's rank correlation, $\rho = 0.73$, $p < 0.0001$. Note a strong increase in co-mapping after fear conditioning specifically for stimuli with larger representational similarity to the CS+ prior to conditioning.

(G) Top: experimental time line of behavioral experiment. During conditioning the same complex sound (CS6) as in Figure 6 was paired with the unconditioned stimulus. During the generalization test, mice were exposed to three non-conditioned sound stimuli. Bottom: increase in freezing behavior in a test session for three non-conditioned stimuli (two with high-response correlation to the CS+ (nonCS+ (a, b)) and one with low response correlation to the CS+ (nonCS+ (c)). Freezing to nonCS+ (c) is not different to time periods without the presentation of a stimulus (blank). Gray lines depict behavior of individual animals and the black line is the mean ± SEM of all animals. *One-way ANOVA with correction for multiple comparisons, first 30 s versus nonCS+ (a, b, c) and blank: $p < 0.05$, nonCS+ (a, b) versus nonCS+ (c) and blank: $p < 0.0001$.

capture structure in massive parallel recordings, changes in population activity are often described using a single, albeit rather abstract metric, such as decoding power (e.g., Figure 3D). Here, the identification of response modes based on the non-linear activation of distinct neuronal subgroups provided a description of population activity at an intermediate and biologically more interpretable level. A response mode represents the mapping of a set of sensory stimuli to the stereotypical activation of a specific subgroup of neurons and thus reflects non-linear properties of auditory perception (Liberman et al., 1967). Using this framework, we found that auditory representations undergo substantial recombination even under environmentally and behaviorally

stable conditions. We observed nearly stationary dynamics involving the gain and loss of response modes as well as changes in the mapping of stimuli to individual response modes.

What are the driving forces underlying the recombination of sensory representations during basal conditions? As the pattern of connectivity is considered a major determinant for the patterns of activity that can arise in neuronal networks, it appears plausible that ongoing remodeling of synaptic connections could underlie the plasticity we observed in our experiments. Indeed, such basal dynamics in connectivity are observed in the mouse auditory cortex during behaviorally stable conditions without need for adaption (Loewenstein et al., 2011, 2015). Interestingly, such spontaneous dynamics in synaptic connections even persist during pharmacological blockade of neuronal activity (Dvorkin and Ziv, 2016; Nagaoka et al., 2016; Rubinski and Ziv, 2015; Yasumatsu et al., 2008), and thus appear to represent a fundamental feature of neuronal circuits. More recently, theoretical modeling has been used to investigate how ongoing synaptic plasticity, as it is observed during basal conditions, affects the long-term stability of activity patterns in a network (Humble et al., 2019; Kappel et al., 2015, 2018; Mongillo et al., 2018; Susman et al., 2019).

For successful navigation in a complex and changing environment, it is essential to balance the discrimination and generalization of sensory stimuli. Generalization and transfer of learning to stimuli sharing perceptual features can be useful, given that natural stimuli typically reoccur in a similar, but rarely in the exact same form. However, excessive stimulus generalization is believed to play an important role in clinical contexts, such as post-traumatic stress disorder (Besnard and Sahay, 2016). In the laboratory setting, stimulus generalization is typically investigated in conditioning paradigms employing simple stimuli, such as pure tones in the auditory domain that can be changed in a unidimensional manner to control their perceptual similarity to the CS+. Predicting the perceptual similarity for naturalistic stimuli that are different from the CS+ along multiple sensory dimensions is challenging.

In our study, we used the degree of representational similarity between the CS+ and non-conditioned stimuli prior to conditioning as a proxy for perceptual similarity, allowing us to infer the extent of learning-induced changes in the sound-evoked activity patterns. This is in line with previous findings in which the representational similarity at the level of the auditory cortex was used as a neurometric measure to predict behavioral categorization of stimuli in a discrimination task (Bathellier et al., 2012). Stimulus generalization is believed to entail a linkage between the CS+ and nonCS+ sharing perceptual features; however, the circuit mechanisms mediating this linkage are still poorly understood (Boddez et al., 2021; Dunsmoor and Paz, 2015). When studying the effect of ACFC on the intrinsic dynamics of auditory representations, we observed an increased co-mapping of non-conditioned stimuli to the same subgroups of neurons that are driven by the CS+, consistent with a contribution of the auditory cortex to stimulus generalization (Aizenberg and Geffen, 2013; Armony et al., 1997; Thompson, 1962). Increased co-activation of shared subgroups of neurons by different stimuli has been suggested to represent a neuronal correlate of the formation of an association between the stimuli. Increased co-activation of shared neuronal subgroups has been reported previously during the formation of an association between the CS+ and the unconditioned stimulus at the level of the amygdala (Grewe et al., 2017) as well as in hippocampal ensembles representing environmental context where both the degree of co-activation and the transfer of fear between conditioned and non-conditioned contexts was dependent on the time elapsed between both experiences (Cai et al., 2016). Together, our findings suggest that generalization follows and additionally reinforces the structure of internal representations.

Of note, our study did not reveal prominent conditioning-induced effects on the neuronal representation of the CS+ that outlasted several days and could be interpreted as a representational expansion in the auditory cortex (Figure 6G). This is in contrast to an earlier study in which guinea pigs were extensively conditioned to pure tones (Weinberger et al., 1993), but consistent with a more recent study in which discriminatory fear conditioning led to an increased representational distance between the paired CS+ and an unpaired CS− in the absence of pronounced response potentiation (Dalmay et al., 2019).

Interestingly, the recombination of representations, i.e., their association as well as their dissociation, ensues to a substantial degree also during basal conditions and remains a matter of research to investigate their functional relevance (Chambers and Rumpel, 2017). We speculate that the ongoing dynamics of subgroups of neurons could support cognitive processes that occur without explicit mental engagement, such as the spontaneous creation of associations (Wallas, 1926) or their forgetting (Richards and Frankland, 2017).

Limitations of the study

Here, we used calcium imaging to record sound-evoked patterns of activity in the mouse auditory cortex. As such, it remains an indirect measure of neuronal activity and offers only limited temporal resolution. It therefore remains unanswered in how far the non-linearly activated subgroups of neurons corresponding to various response modes reflect Hebbian cell assemblies. A better temporal resolution of the fine structure of activity patterns as much as a reconstruction of the underlying connectivity will be needed to resolve this question in the future. Furthermore, we used Pearson correlation of sound-evoked activity patterns as a measure of representational similarity that provided a good, albeit not perfect, predictor of the effects observed following conditioning. This indicates that another measure could provide an even better metric for representational similarity.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mouse maintenance
- METHODS DETAILS
  - Molecular cloning

- rAAV production
- Stereotaxic injections
- Cranial window implantation
- Habituation to awake chronic two-photon imaging
- Sound presentation
- *In vivo* two photon imaging
- Auditory cued fear conditioning
- Conditioning session
- Memory and generalization test session
- Quantitative analysis of freezing behavior
- Confocal imaging
- Image processing of chronic *in vivo* two-photon data
- ROI inclusion criteria
- Calculation of $\Delta F/F_0$ and deconvolution
- Stimulus-evoked sound responsiveness of single cells
- Sound response profiles of single cells
- FOV inclusion criteria
- Definition of response modes
- Mode-associated responsiveness of single cells
- Pairwise sound decoding based on full response vectors
- Correlation of stimulus responses
- Difference of stimulus responsiveness on day one and day seven
- Difference of stimulus co-mapping to the CS + on day seven between the two experimental cohorts
- Sound decoding based on a 34-fold classifier
- Dependence of decoding on number of modes and stimuli per mode

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Freezing behavior
  - Comparison of neuronal population dynamics between basal and fear conditioning groups
  - Difference of stimulus responsiveness and co-mapping to the CS+
  - Comparison of cellular sound response profiles after sorting with data from different sessions or trials
  - Decoding comparison
  - Comparison of mean correlations of stimulus response vectors between the experimental cohorts

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110340.

ACKNOWLEDGMENTS

We thank Drs. Yonatan Loewenstein, Gianluigi Mongillo, and members of the Kaschube and Rumpel labs for helpful discussions and Dr. Noam Ziv for critical reading of an earlier version of the manuscript. This work was supported by research grant DFG CRC1080-C05 (to S.R.), DFG/ANR project no. 431393205 (to S.R.), DFG SPP 2041 project no. 347573108 (to S.R. and M.K.), and DFG DIP "Neurobiology of Forgetting" (to S.R. and M.K.).

AUTHOR CONTRIBUTIONS

Conceptualization, M.K. and S.R.; methodology, D.F.A., J.-B.E., and C.P.; software, J.B.E.; formal analysis, D.F.A., J.-B.E., L.E., and A.R.C.; investigation, D.F.A.; writing – original draft, D.F.A., J.-B.E., M.K., and S.R.; writing – review & editing, L.E. and A.R.C.; supervision, M.K. and S.R.; funding acquisition, M.K. and S.R.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 21, 2021
Revised: November 16, 2021
Accepted: January 14, 2022
Published: February 8, 2022

REFERENCES

Aizenberg, M., and Geffen, M.N. (2013). Bidirectional effects of aversive learning on perceptual acuity are mediated by the sensory cortex. Nat. Neurosci. *16*, 994–996.

Armony, J.L., Servan-Schreiber, D., Romanski, L.M., Cohen, J.D., and LeDoux, J.E. (1997). Stimulus generalization of fear responses: effects of auditory cortex lesions in a computational model and in rats. Cereb. Cortex *7*, 157–165.

Aschauer, D., and Rumpel, S. (2018). The sensory neocortex and associative memory. Curr. Top. Behav. Neurosci. *37*, 177–211.

Atencio, C.A., and Schreiner, C.E. (2013). Auditory cortical local subnetworks are characterized by sharply synchronous activity. J. Neurosci. *33*, 18503–18514.

Bathellier, B., Ushakova, L., and Rumpel, S. (2012). Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron *76*, 435–449.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate—a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B *57*, 289–300.

Besnard, A., and Sahay, A. (2016). Adult hippocampal neurogenesis, fear generalization, and stress. Neuropsychopharmacology *41*, 24–44.

Boddez, Y., Finn, M., and De Houwer, J. (2021). The (shared) features of fear: toward the source of human fear responding. Curr. Opin. Psychol. *41*, 113–117.

Cai, D.J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., Wei, B., Veshkini, M., La-Vu, M., Lou, J., et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. Nature *534*, 115–118.

Chambers, A.R., and Rumpel, S. (2017). A stable brain from unstable components: emerging concepts and implications for neural computation. Neuroscience *357*, 172–184.

Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature *499*, 295–300.

Cho, J.H., Rendall, S.D., and Gray, J.M. (2017). Brain-wide maps of Fos expression during fear learning and recall. Learn. Mem. *24*, 169–181.

Clopath, C., Bonhoeffer, T., Hubener, M., and Rose, T. (2017). Variance and invariance of neuronal long-term representations. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20160161.

Dalmay, T., Abs, E., Poorthuis, R.B., Hartung, J., Pu, D.L., Onasch, S., Lozano, Y.R., Signoret-Genest, J., Tovote, P., Gjorgjieva, J., et al. (2019). A critical role for neocortical processing of threat memory. Neuron *104*, 1180–1194.e7.

Deitch, D., Rubin, A., and Ziv, Y. (2021). Representational drift in the mouse visual cortex. Curr. Biol. *31*, 4327–4339.e6.

Driscoll, L.N., Pettit, N.L., Minderer, M., Chettih, S.N., and Harvey, C.D. (2017). Dynamic reorganization of neuronal activity patterns in parietal cortex. Cell *170*, 986–999.e16.

Dunsmoor, J.E., and Paz, R. (2015). Fear generalization and anxiety: behavioral and neural mechanisms. Biol. Psychiatry *78*, 336–343.

Dvorkin, R., and Ziv, N.E. (2016). Relative contributions of specific activity histories and spontaneous processes to size remodeling of glutamatergic synapses. PLoS Biol. *14*, e1002572.

Gillet, S.N., Kato, H.K., Justen, M.A., Lai, M., and Isaacson, J.S. (2017). Fear learning regulates cortical sensory representations by suppressing habituation. Front. Neural Circuits *11*, 112.

Grewe, B.F., Grundemann, J., Kitch, L.J., Lecoq, J.A., Parker, J.G., Marshall, J.D., Larkin, M.C., Jercog, P.E., Grenier, F., Li, J.Z., et al. (2017). Neural ensemble dynamics underlying a long-term associative memory. Nature *543*, 670–675.

Grimm, D., Kay, M.A., and Kleinschmidt, J.A. (2003). Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. Mol. Ther. *7*, 839–850.

Hainmueller, T., and Bartos, M. (2018). Parallel emergence of stable and dynamic memory engrams in the hippocampus. Nature *558*, 292–296.

Huber, D., Gutnisky, D.A., Peron, S., O'Connor, D.H., Wiegert, J.S., Tian, L., Oertner, T.G., Looger, L.L., and Svoboda, K. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. Nature *484*, 473–478.

Humble, J., Hiratsuka, K., Kasai, H., and Toyoizumi, T. (2019). Intrinsic spine dynamics are critical for recurrent network learning in models with and without autism spectrum disorder. Front. Comput. Neurosci. *13*, 38.

Kappel, D., Habenschuss, S., Legenstein, R., and Maass, W. (2015). Network plasticity as Bayesian inference. PLoS Comput. Biol. *11*, e1004485.

Kappel, D., Legenstein, R., Habenschuss, S., Hsieh, M., and Maass, W. (2018). A dynamic connectome supports the emergence of stable computational function of neural circuits through reward-based learning. eNeuro *5*. https://doi.org/10.1523/ENEURO.0301-17.2018.

Kato, H.K., Gillet, S.N., and Isaacson, J.S. (2015). Flexible sensory representations in auditory cortex driven by behavioral relevance. Neuron *88*, 1027–1039.

Kopec, C.D., Kessels, H.W., Bush, D.E., Cain, C.K., LeDoux, J.E., and Malinow, R. (2007). A robust automated method to analyze rodent motion during fear conditioning. Neuropharmacology *52*, 228–233.

Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4.

Lai, C.S.W., Adler, A., and Gan, W.B. (2018). Fear extinction reverses dendritic spine formation induced by fear conditioning in the mouse auditory cortex. Proc. Natl. Acad. Sci. U S A *115*, 9306–9311.

Letzkus, J.J., Wolff, S.B., Meyer, E.M., Tovote, P., Courtin, J., Herry, C., and Luthi, A. (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. Nature *480*, 331–335.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., and Studdert-Kennedy, M. (1967). Perception of the speech code. Psychol. Rev. *74*, 431–461.

Loewenstein, Y., Kuras, A., and Rumpel, S. (2011). Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. J. Neurosci. *31*, 9481–9488.

Loewenstein, Y., Yanover, U., and Rumpel, S. (2015). Predicting the dynamics of network connectivity in the neocortex. J. Neurosci. *35*, 12535–12544.

Mankin, E.A., Sparks, F.T., Slayyeh, B., Sutherland, R.J., Leutgeb, S., and Leutgeb, J.K. (2012). Neuronal code for extended time in the hippocampus. Proc. Natl. Acad. Sci. U S A *109*, 19462–19467.

Margolis, D.J., Lutcke, H., Schulz, K., Haiss, F., Weber, B., Kugler, S., Hasan, M.T., and Helmchen, F. (2012). Reorganization of cortical population activity imaged throughout long-term sensory deprivation. Nat. Neurosci. *15*, 1539–1546.

Moczulska, K.E., Tinter-Thiede, J., Peter, M., Ushakova, L., Wernle, T., Bathellier, B., and Rumpel, S. (2013). Dynamics of dendritic spines in the mouse auditory cortex during memory formation and memory recall. Proc. Natl. Acad. Sci. U S A *110*, 18315–18320.

Mongillo, G., Rumpel, S., and Loewenstein, Y. (2018). Inhibitory connectivity defines the realm of excitatory plasticity. Nat. Neurosci. *21*, 1463–1470.

Nagaoka, A., Takehara, H., Hayashi-Takagi, A., Noguchi, J., Ishii, K., Shirai, F., Yagishita, S., Akagi, T., Ichiki, T., and Kasai, H. (2016). Abnormal intrinsic dynamics of dendritic spines in a fragile X syndrome mouse model in vivo. Sci. Rep. *6*, 26651.

Nathanson, J.L., Yanagawa, Y., Obata, K., and Callaway, E.M. (2009). Preferential labeling of inhibitory and excitatory cortical neurons by endogenous tropism of adeno-associated virus and lentivirus vectors. Neuroscience *161*, 441–450.

Pavlov, I.P., and Anrep, G.V. (1927). Conditioned reflexes; an investigation of the physiological activity of the cerebral cortex. Ann. Neurosci. *17*, 136–141.

Peter, M., Scheuch, H., Burkard, T.R., Tinter, J., Wernle, T., and Rumpel, S. (2012). Induction of immediate early genes in the mouse auditory cortex after auditory cued fear conditioning to complex sounds. Genes Brain Behav. *11*, 314–324.

Peter, M., Aschauer, D.F., Rose, R., Sinning, A., Grossl, F., Kargl, D., Kraitsy, K., Burkard, T.R., Luhmann, H.J., Haubensak, W., et al. (2021). Rapid nucleus-scale reorganization of chromatin in neurons enables transcriptional adaptation for memory consolidation. PLoS One *16*, e0244038.

Quirk, G.J., Armony, J.L., and LeDoux, J.E. (1997). Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron *19*, 613–624.

Richards, B.A., and Frankland, P.W. (2017). The persistence and transience of memory. Neuron *94*, 1071–1084.

Rokni, U., Richardson, A.G., Bizzi, E., and Seung, H.S. (2007). Motor learning with unstable neural representations. Neuron *54*, 653–666.

Rubinski, A., and Ziv, N.E. (2015). Remodeling and tenacity of inhibitory synapses: relationships with network activity and neighboring excitatory synapses. PLoS Comput. Biol. *11*, e1004632.

Rule, M.E., O'Leary, T., and Harvey, C.D. (2019). Causes and consequences of representational drift. Curr. Opin. Neurobiol. *58*, 141–147.

Schoonover, C.E., Ohashi, S.N., Axel, R., and Fink, A.J.P. (2021). Representational drift in primary olfactory cortex. Nature *594*, 541–546.

See, J.Z., Atencio, C.A., Sohal, V.S., and Schreiner, C.E. (2018). Coordinated neuronal ensembles in primary auditory cortical columns. eLife *7*, e35587.

Susman, L., Brenner, N., and Barak, O. (2019). Stable memory with unstable synapses. Nat. Commun. *10*, 4441.

Thompson, R.F. (1962). Role of the cerebral cortex in stimulus generalization. J. Comp. Physiol. Psychol. *55*, 279–287.

Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., and Paninski, L. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. J. Neurophysiol. *104*, 3691–3704.

Wallas, G. (1926). The Art of Thought (Solis Press).

Weinberger, N.M. (2004). Specific long-term memory traces in primary auditory cortex. Nat. Rev. Neurosci. *5*, 279–290.

Weinberger, N.M., Javid, R., and Lepan, B. (1993). Long-term retention of learning-induced receptive-field plasticity in the auditory cortex. Proc. Natl. Acad. Sci. U S A *90*, 2394–2398.

Yang, Y., Liu, D.Q., Huang, W., Deng, J., Sun, Y., Zuo, Y., and Poo, M.M. (2016). Selective synaptic remodeling of amygdalocortical connections associated with fear memory. Nat. Neurosci. *19*, 1348–1355.

Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., and Kasai, H. (2008). Principles of long-term dynamics of dendritic spines. J. Neurosci. *28*, 13592–13608.

Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nat. Neurosci. *16*, 264–266.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
rAAV (ITR-hSyn-GCaMP6m-WPRE-ITR)	This paper	N/A
rAAV (ITR-hSyn-H2B::mCherry-WPRE-ITR)	This paper	N/A
Deposited data		
In vivo two-photon imaging data	This paper	https://doi.org/10.12751/g-node.trwj8c
Experimental models: Organisms/strains		
Mouse: CB57BL/6J	Jackson Laboratory	Strain #000664; RRID: IMSR_JAX:000664
Recombinant DNA		
pAAV-hSyn-hChR2(H134R)-EYFP	Addgene	Addgene 26973; RRID:Addgene_26973
pAAV-hSyn-GCaMP6m-WPRE	This paper	N/A
pAAV-hSyn-H2B::mCherry-WPRE	This paper	N/A
Software and algorithms		
MATLAB R2007a, R2016b, R2019b	MathWorks, Natick, MA, USA	N/A
Image processing of chronic two-photon data	This paper	https://doi.org/10.5281/zenodo.5822486

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Simon Rumpel (sirumpel@uni-mainz.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Original imaging data, corresponding ROI coordinates for extraction of cellular calcium activity and sound stimulation data have been deposited at G-Node (www.g-node.org) and is made publicly available as of the date of publication (G-Node: https://doi.org/10.12751/g-node.trwj8c).

All original image processing code has been deposited at the Zenodo online depository and can be accessed at Zenodo: https://doi.org/10.5281/zenodo.5822486.

Any additional information that may be required to re-analyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse maintenance

Experimental subjects for imaging experiments were male CB57BL/6J mice of eight to twelve weeks of age from Jackson laboratory (strain #000664). For behavior experiments investigating the generalization of different auditory stimuli, experimental subjects were male C57BL/6JRj of eight weeks of age from Janvier laboratories. Before surgical procedures, mice were kept in groups of five, and housed in 530 cm^2 cages on a 12h light/dark cycle with unlimited access to dry food and water. Experiments were carried out during the light period. All animal experiments were performed in accordance with the Austrian laboratory animal law guidelines for animal research and had been approved by the Viennese Magistratsabteilung 58 (Approval M58/00236/2010/6) and the Landesuntersuchungsamt Rheinland Pfalz (Approval 23 177-07/G 17-1-051).

METHODS DETAILS

Molecular cloning

For the generation of a recombinant AAV (rAAV) genome encoding for GCaMP6m under the human SynapsinI promoter (*phSyn*), a plasmid containing the inverted terminal repeats (ITRs) of AAV, *phSyn* (Addgene plasmid 26973), Woodchuck Hepatitis Prottranscriptional Regulatory Element (WPRE), and a human Growth Hormone polyadenylation site (*hGH-pA* site) was digested using *BamHI* and *AccIII* and the gene coding for GCaMP6m was PCR amplified from a commercially available plasmid (Addgene plasmid 40754) and inserted. Finally, the plasmid was digested with *AccIII* and *HindIII* to excise the original transgene and 3′overhangs were blunted and 5′overhangs were filled in using Klenow fragment.

For the generation of a recombinant AAV genome encoding for H2B-mCherry fusion protein under the *phSyn*, a gene coding for mCherry was PCR amplified and inserted into a plasmid containing a gene for H2B directly after its coding sequence using *ClaI* and *SpeI* to produce a fusion gene. The H2B-mCherry fusion gene was PCR amplified and inserted into a plasmid containing ITRs, *phSyn*, *WPRE*, and *hGH-pA* using *KpnI* and *HindIII*. Finally, the *WPRE* was removed using *HindIII* and *XhoI,* and 3′overhangs were blunted and 5′overhangs were filled in using Klenow fragment.

rAAV production

All rAAV vectors described were produced in HEK 293 cells by using a helper virus free, two-plasmid based production method (Grimm et al., 2003) based on a commercially available AAV helper free system (Agilent Technologies, CA, USA; catalog# 240,071). Briefly, *HEK293* cells were transfected by using the calcium phosphate method. 72 h post transfection, cells were harvested and collected by centrifugation (2,500 x g, 20 min at 4°C). Cell pellets were resuspended in resuspension buffer and lysed by three consecutive freeze/thaw cycles. For removal of genomic DNA, cell lysates were incubated with benzonase (50 U/mL) for one hour at 37°C. Subsequently, rAAV particles were precipitated with CaCl2 (25 mM) followed by PEG precipitation (8% PEG-8000, 500 mM NaCl). After resuspension of PEG precipitates in 50 mM HEPES, 150 mM NaCl, 25 mM EDTA, pH 7.4 overnight at 4°C, rAAV particles were further purified by CsCl density gradient centrifugation. Fractions from CsCl density gradients were analyzed by measuring the refractory index. Samples within a refractory index ranging from 1.3774 to 1.3696 were pooled and dialyzed against PBS for removal of CsCl by using dialysis cassettes with a molecular weight cutoff of 20 kDa (Thermo Scientific, MA, USA; catalog# 87738). Finally, rAAV preparations were concentrated by using ultrafiltration units with a molecular weight cutoff of 50 kDa (Millipore, MA, USA; catalog# UFC905024). After addition of glycerol to a final concentration of 10%, rAAV preparations were sterile filtered with Millex-GV filter units (Millipore, MA, USA; catalog# SLGV013SL), frozen in liquid nitrogen, and subsequently stored in aliquots at −80°C. Genomic titers of purified rAAV stocks were determined by isolation of viral DNA (Viral Xpress DNA/RNA Extraction Reagent, Millipore, MA, USA; catalog# 3095) and subsequent qPCR analysis using primers specific for *phSyn*.

Stereotaxic injections

All surgical equipment was sterilized with 70% v/v ethanol before use. Animals were deeply anesthetized with a mixture of ketamine and medetomidine (KM; 2.5 mg ketamine-HCl and 0.02 mg medetomidine-HCl/25 g mouse weight) injected intraperitoneally, and positioned in a stereotaxic frame (Kopf Instruments, Tujunga, CA, USA; Stereotaxic System Kopf, 1900). The eyes were protected from dehydration and intensive light exposure using sterile eye gel (Alcon Pharma, Novartis, CHE; Thilo-Tears Gel) and a piece of aluminum foil. Lidocaine was applied as local anesthetic subcutaneously before exposure of the skull. The scalp was washed with a 70% v/v ethanol in water solution and a cut along the midline revealed the skull. A small hole was drilled into the skull above the auditory cortex using a stereotaxic motorized drill (Kopf Instruments, Tujunga, CA, USA; Model, 1911 Stereotaxic Drilling Unit) leaving the dura mater intact. Injections were performed perpendicular to the surface of the skull. Virus solution consisted of a mixture of two different recombinant AAV viruses (rAAV2/8 ITR-phSyn-GCaMP6m-WPRE-hGHpolyA-ITR; titer: $1.75 * 10^{11}$ viral genomes (VG)/ml; rAAV2/8 ITR-phSyn-H2BmCherry-hGHpolyA-ITR; titer: $2 * 10^{13}$ VG/ml) in PBS. The virus mixture was loaded into a thin glass pipette and 150 nL were injected at a flow rate of 20 nL/min (World Precision Instruments, Sarasota, FL, USA; Nanoliter, 2000 Injector) in five locations along the anterior-posterior axis, resulting in a total injection volume of 750 nL. Stereotactic coordinates were: 4.4, −2,5/-2.75/-3/-3.25/-3.5, 2.5 (in mm, caudal, lateral, and ventral in reference to bregma). Glass pipettes (World Precision Instruments, Sarasota, FL, USA; Glass Capillaries for Nanoliter, 2000; Order# 4878) had been pulled with a long taper and the tip was cut to a diameter of 20–40 μm. After the injection, the pipette was left in place for three minutes, before being slowly withdrawn and moved to the next coordinate. After completion of the injection protocol, the skin wound was sealed using tissue adhesive (3M Animal Care Products, St. Paul, MN, USA; 3M Vetbond Tissue Adhesive), and anesthesia was neutralized with 0.02 mL atipamezole. Mice were monitored daily and intraperitoneal injections of carprofen (0.2 mL of 0.5 mg/mL stock) were applied on the first days after surgery.

Cranial window implantation

Two weeks after stereotactic injections, animals were anesthetized using isoflurane (Abbott Animal Health, IL, USA; IsoFlo). All surgical equipment and glass coverslip were sterilized with 70% v/v ethanol before use. Anesthesia was initialized in a glass desiccator filled with an isoflurane/air mixture. Anesthetized animals were mounted on a stereotaxic frame (Kopf Instruments, Tujunga, CA, USA; Stereotaxic System Kopf, 1900) and the head was positioned using ear, teeth, and a custom-made v-shaped head holder. Anesthesia was maintained by delivery of a 1.5 to 2.4% isoflurane/air mixture with a vaporizer (High Precision Instruments, MT; Univentor

400 Anesthesia Unit) at a flow rate of around 200 mL/min to the snout. 0.02 mL dexamethasone (4 mg/mL) was administered intramuscularly to the quadriceps, as well as 0.02 mL carprofen (0.5 mg/mL) intraperitoneally. The eyes were protected from dehydration and intensive light exposure using sterile eye gel (Alcon Pharma, Novartis, CHE; Thilo-Tears Gel) and a piece of aluminum foil. A local anesthetic (lidocaine/epinephrine (Gebro Pharma, Austria)) was applied subcutaneously before exposure of the skull. The scalp was washed with a 70% v/v ethanol in water solution and a flap of skin covering temporal, both parietal regions and part of the occipital bone was removed. The musculus temporalis was injected with lidocaine/epinephrine (Gebro Pharma, Austria) as an additional anesthetic and to minimize bleeding. Subsequently, the muscle was partly removed with a surgical scalpel and forceps to expose the right temporal bone. Using a fine motorized drill, the bones of the skull were smoothened, and part of the zygomatic process was removed. The surface was cleaned using cortex buffer and a 2% v/v hydrogen peroxide in water solution, and covered with a thin layer of one component-instant glue (Carl Roth, Germany; Roti coll). A thin layer of dental cement (Lang Dental, IL, USA; Ortho-Jet) was applied onto the skull, sparing the area of the temporal bone above the auditory cortex. A rectangular groove of about 2 mm by 3 mm was carefully drilled into the skull above the auditory cortex, and the bone was carefully lifted using scalpel and forceps. The exposed area was carefully cleaned and kept moist using sterile sponge (Pfizer, NY, USA; Gelfoam) and cortex buffer. The craniotomy was covered with a small circular cover glass (Electron Microscopy Sciences, PA, USA; five mm diameter, catalog# 72,195-05), and sealed with 1.2% low-melting agarose (Sigma Aldrich, MO, USA; Agarose Type IIIA). The cover glass was finally set in place with one component-instant glue and dental cement. In order to position the animal under the microscope with the objective facing the window plane perpendicularly, a custom-made titanium head post was mounted on the implant above the window and embedded with dental cement. After dental cement had dried, animals were placed back in a pre-warmed cage. After the surgical procedure, animals recovered for at least one week before further handling.

Habituation to awake chronic two-photon imaging

Animals were habituated to handling at the two-photon microscope. Therefore, animals were mildly water deprived and fixated under the objective in a custom-made acrylic glass tube, using a custom-made head post implant. The mouse head was laterally tilted such that the surface of the auditory cortex aligns approximately with the horizontal plane. During habituation, head fixation lasted for a minimum of 30 min each day, and animals were given access to a 5% m/v sucrose in water solution. This was repeated for at least five days until animals accommodated to the head fixation apparatus, showed reduced signs of stress and less body movements (typically consisting of few second long running bouts). The full sound stimulus set later used for recording of sound-evoked activity, was repeatedly presented. Hence, animal subjects did extensively experience all sensory stimuli before any data acquisition.

Sound presentation

All sounds were delivered free field at 192 kHz sampling rate in a soundproof booth by a custom-made system consisting of a linear amplifier and a ribbon loudspeaker (Audiocomm, Austria) placed in 25 cm distance to the mouse head. The transfer function between the loudspeaker and the location of the mouse ear was measured using a probe microphone (Brüel&Kjær, Bremen, Germany; 4939-L-002) and compensated numerically by filtering the sound files with the inverse transfer function to obtain a flat frequency response at the mouse ear (between 0.5 kHz and 64 kHz ±4 dB). Sound control and equalization was performed by a custom MATLAB program running on a standard personal computer equipped with a Lynx 22 sound card (Lynx Studio Technology, CA, USA). The stimulus set consisted of 34 sound stimuli (19 pure-tone pips (50ms; 2 to 45 kHz separated by a quarter octave) and 15 complex sounds (70ms)) separated by one-second-intervals and played at 80 dB sound pressure level. The complex sounds in the stimulus set were characterized by broad frequency content and temporal modulations, generated from arbitrary samples of music pieces or animal calls replayed at fourfold speed. All stimulus on- and offsets were smoothened with a ten-ms-long half-period cosine function.

In vivo two photon imaging

The two-photon microscope (Prairie Technologies, WI, USA; Ultima IV) was comprised of a 20×-objective (Olympus, Tokyo, Japan; XLUMPlan FI, NA = 0.95) and a pulsed laser (Coherent, CA, USA; Chameleon Ultra). Both fluorophores (GCaMP6m and mCherry) were co-excited at 920 nm wavelength, and separated by emission using a fluorescence filter cube (filter one: BP 480-550 nm; filter two: LP 590 nm; dichromatic mirror: DM 570 nm; Olympus, Tokyo, Japan; U-MSWG2). Full frame imaging was performed using a field of view of 367 × 367 μm (pixel size: 256x128) and images were acquired at five Hz frame rate (sampling period: 196,86ms). In the last habituation session, several field of views (FOVs) at different xy-positions in layer 2/3 (about 150–300 μm depth from cortical surface) were screened for the presence of reliable sound responses. FOVs with reliable sound responses were repeatedly imaged at a two-day interval, using the stimulus set described above. Each stimulus was presented for at least 20 repetitions per FOV in pseudo-randomized order. Next, the focal plane was moved 50 μm in the z axis and data was acquired for a second FOV with the same xy-coordinates. Between imaging periods, animals were given access to few drops of a 5% w/v sucrose in water solution.

Auditory cued fear conditioning

The behavioral setup for imaging experiments was controlled by a personal computer with WINDOWS XP Professional, Version 2002, SP2 (Microsoft, Redmond, WA, USA) operating system running custom MATLAB R2007a software (MathWorks, Natick, MA, USA). The behavioral setup for experiments investigating the generalization of different auditory stimuli was controlled by a personal computer with WINDOWS 10, Version 2021, (Microsoft, Redmond, WA, USA) operating system running custom MATLAB R2016b

software (MathWorks, Natick, MA, USA). All behavioral experiments were performed in an isolation cubicle (H10-24, Coulbourn Instruments, Whitehall, PA, USA) which was equipped with white LEDs as house light, a microphone and a CCD KB-R3138 camera with infrared LEDs (LG Electronics Austria, Vienna, Austria) which was connected to a Cronos frame grabber (Matrox, Dorval, QC, Canada). The conditioning chamber (25 × 25 × 42 cm, model H10-11M-TC, modified, Coulbourn Instruments) was combined either with a stainless-steel shock floor or a grid floor. A custom-made cartridge (round or quadrangular) formed the walls of the chamber in order to create different local environmental contexts. Foot shocks were delivered via an external shocker (Precision Animal shocker, Coulbourn Instruments). Sounds were played from an L-22 soundcard with a maximal sampling frequency of 192 kHz (Lynx Studio Technology, Costa Mesa, CA, USA) and delivered via an amplifier (Model SLA-1, Applied Research and Technology, TEAC Europe GmbH, TASCAM Division, Wiesbaden, Germany), a modified equalizer (Model #351, Applied Research and Technology, TEAC Europe GmbH, TASCAM Division, Wiesbaden, Germany) and a custom-made speaker for free field delivery of sounds. The sound stimuli used were from the stimulus set used for *in vivo* two-photon calcium imaging. 70 ms stimuli were repeated 15 times with a one-second-interval, resulting in a total duration of 14.07 s. On- and offsets of stimuli were smoothed with a 10-millisecond long half-period cosine function. Sound levels for all stimuli used were normalized to a mean power of 78 dB sound pressure level (SPL). Peak sound levels ranged from 83 to 89 dB SPL.

Conditioning session

In the conditioning environment, lights were turned on (~20-30 lux), and roundish cartridges were used as walls of the chamber. A mild residual ethanol odor was present from previous cleaning of the chamber. Mice were placed in the chamber directly before the start of each session. After at least 1 min baseline (60–90 s) five sound-shock pairings (0.75 mA, one second, immediately following the sound) were given with a randomized inter-stimulus-interval ranging from 50 to 75 s (paired).

Memory and generalization test session

Four days after the conditioning session (i.e. one day after the two-photon imaging paradigm was completed), mice were tested for a conditioned freezing response. In order to create a different environmental context, quadrangular cartridges were used as chamber walls, lights were turned off, and home cage bedding was placed underneath the metal grid to provide a familiar odor. In the memory test session for animals in the imaging experiments (Figure 6B), after at least one minute of baseline (60–90 s), the conditioned stimulus (CS+) and one non-conditioned sound stimulus (nonCS+ (a))were presented in three randomized presentation blocks with an inter-stimulus-interval of twenty to thirty seconds. In the generalization test session for animals in the experiments investigating the generalization of different auditory stimuli (Figures 7G, S12A and S12C), after at least one minute of baseline (60–90 s), three non-conditioned sound stimuli were presented in three randomized presentation blocks with an inter-stimulus-interval of twenty to thirty seconds. In the memory test session for animals in the experiments investigating naive freezing behavior to different auditory stimuli (Figure S12B), after at least one minute of baseline (60–90 s), five non-conditioned sound stimuli were presented in three randomized presentation blocks with an inter-stimulus-interval of twenty to thirty seconds.

Quantitative analysis of freezing behavior

During conditioning, memory and generalization test session, movies were recorded at a frame rate of 2.8 frames per seconds. Movies were analyzed offline based on a similar approach as described previously (Kopec et al., 2007), which provides a rapid and unbiased analysis of animal behavior. In short, the number of 'significant motion pixels' (SMP), i.e. pixels which varied by a fixed threshold of gray values, was calculated for all pairs of consecutive frames using a custom MATLAB R2007a or R2019b script (MathWorks). For each movie, the size of the mouse was estimated by the median SMP value of the 25% highest SMPs calculated from pairs of frames which were recorded at least two minutes apart – thus capturing the mouse likely at different positions in the chamber. The threshold for freezing was defined as fewer SMP than corresponding to 0.3% of the mouse size, which separates SMP values during freezing and movement periods. Baseline freezing was assessed during 60–90 s baseline period of each protocol run.

Confocal imaging

Mice were deeply anesthetized and perfused with a PBS/Heparin solution and subsequently with a 4%PFA solution following standard procedures. Brain sections of 70 μm thickness were cut on a vibratome (Leica Biosystems, Germany; VT-1000). Next, they were incubated for 30 min in a 5 mg/L 4′, 6-diamidino-2-phenylindole (DAPI) solution, and mounted on cover slips. Confocal images were acquired on a LSM780 microscope (Carl Zeiss, Germany) using a 40× immersion objective (Objective Plan-Apochromat 40x/1.4 Oil DIC M27, Carl Zeiss, Germany).

Image processing of chronic *in vivo* two-photon data

In order to track cells across days, the optimal affine transformation was identified to register regions of interest (ROI), encompassing the soma of individual neurons, onto each frame of the time series recorded from the same FOV across several days. ROIs were selected independently by two human experts and can be described by a set of several hundred points marking the centers of the mostly spherical neuronal somata. This set of points was transformed for each frame by a two-dimensional affine transformation. The objective function value for the optimization of this transformation is the pixel-wise overlap between a band-pass filtered and binarized image of each frame and a mask generated from the transformed ROIs by drawing a circle with a three-pixel (4.30 μm)

radius around the center of each ROI. This six-dimensional optimization problem was solved numerically using MATLAB's implementation of the Nelder-Mead-Simplex algorithm (fminsearch). This was done in two iterations, first for the entire frame, then for four equally sized horizontal segments to correct for full frame movements during the two-photon microscope scanning. In a third iteration individual ROIs were moved to the maximum in a two-pixel (2.87 μm) surrounding of a low-pass filtered image to allow for slight local distortions.

ROI inclusion criteria

Four quality criteria were defined in order to only include cells in the analysis that had a reliably present signal in the H2B:mCherry channel marking the neuronal somata. This was done on a frame-by-frame basis, so that at each given time point a cell was either reliably present or excluded. The number of cells excluded in each of these steps are shown in Figure S2.

Nearest Neighbor Distance (NND)

Strongly overlapping cells in a given frame, i.e. cells with a center-to-center distance below three pixels (4.30 μm), were defined as unreliable in that respective frame. Thus, the chance to wrongly label individual cells was minimized.

Normalized Soma Signal Intensity (NSSI)

For each cell at each time point, the difference between the mean signal intensity in the soma (two-pixel radius; 2.87 μm) and the mode of the intensity of the surrounding (ten-pixel radius; 14.34 μm) was computed and normalized by the 95-percentile of this difference. Cells with an intensity close to the background, an NSSI below the value of 0.2 were excluded.

Objective Function Value (OFV)

The optimization described above resulted in the alignment and an objective function value, which describes the pixel wise overlap of the frame and the template. In order to rule out movement artifacts, individual frames in which the OFV was less than three standard deviations below the mode of the OFV for a given FOV were rejected.

Soma Signal to Noise ratio (SSN)

The difference of the mean intensity of the soma (two-pixel radius; 2.87 μm) and the mode of the intensity of the surrounding (ten-pixel radius; 14.34 μm) was defined as signal. The standard deviation of a jittered version of the signal (same radii, but pseudo-random location of the "soma" in the ten-pixel radius) was defined as noise. In order to be included in the analysis, cells had to have an SSN value above one.

All quality criteria were tested and cells were excluded on a frame-by-frame basis. Excluded time points were treated as missing entries in the data. Cells that were not reliably detected on at least ten trials for each stimulus on a given day were completely excluded from the analysis.

Calculation of $\Delta F/F_0$ and deconvolution

The baseline F_0 used to compute the $\Delta F/F_0$ was defined as a moving rank order filter, the 30th percentile of the 200 surrounding frames (100 before and 100 after). This $\Delta F/F_0$ was then deconvolved using the algorithm published by Vogelstein and colleagues (Vogelstein et al., 2010).

Stimulus-evoked sound responsiveness of single cells

To classify single cells as sound responsive or not, all trials from a given stimulus were compared in a rank-sum test against twenty randomly picked patterns of spontaneous activity (from periods without sound presentation). A cell was classified as significantly responsive, if the p value was below 0.01 after a Benjamini-Hochberg correction for multiple comparisons against number of days (4), number of stimuli (34), and number of cells (21506) for at least one stimulus (Benjamini and Hochberg, 1995).

Sound response profiles of single cells

For each significantly sound responsive cell, sound response vectors to pure-tone frequencies and complex sound stimuli were max-normalized to the stimulus with highest response amplitude on the given day. The selection of cells for each day and sorting of cells on the y axis was performed either on significantly sound responsive cells and their tuning from day one (Figure 2A), the given day (Figure 2C), or from day seven (Figure 2E).

For the analysis to control for a sampling bias, the sorting of cells on the y axis was performed with a subset of either the first or the second half of trials of a session (Figure S3A).

FOV inclusion criteria

We included FOVs in our analysis that satisfied the following three criteria: (A) FOVs needed to contain at least 100 ROIs (i.e. neurons): which fulfilled the quality criteria described above, (B) FOVs needed more than ten significantly sound responsive neurons on each day and (C) neuronal populations in the FOVs needed to respond to at least four stimuli on at least one day.

Definition of response modes

Response modes were defined for a given FOV. For each trial *i*, the population response of *n* simultaneously recorded neurons was characterized by an *n*-dimensional vector v. Each entry of v was the mean deconvolved activity recorded in a time bin of 400 ms after stimulus onset. The response similarity between two stimuli p and q was then determined by

$$S(p,q) = \frac{1}{M_p M_q} \sum_{i=1}^{M_p} \sum_{j=1}^{M_q} \rho(v_{p,i} v_{q,j})$$

with trial numbers M_p, M_q, and Pearson's correlation coefficient $\rho(v_1,v_2)$. Note that stimuli p and q may refer to stimulus pairs presented at the same day (Figure 4) or at two different days (Figures 5 and 7), in which case the same stimulus presented at different days is formally treated as two different stimuli. Response reliability for a given stimulus, at a given day was assessed by *S(p,p)* (the mean correlation over all pairs of trials, excluding pairs with i = j, i.e. same trials). The response to a stimulus, for which $S(p,p) > 0.4$ was deemed reliable. Response modes were then estimated by hierarchical clustering of response similarity *S(p,q)*, restricted to (day-specific) stimuli p with reliable response, with 1-*S(p,q)* as metric and unweighted average linkage clustering as linkage criteria. All responses of stimuli with non-reliable response ($S(p,p) < 0.4$) were assigned to the "null-mode" (0-mode). Choosing this threshold we found that most prominent clusters were well captured, as we verified by visual inspection of all FOVs. Importantly, our overall results are qualitatively similar when the threshold for response reliability was set to 0.3 or 0.5 (data not shown). The hierarchical clustering algorithm provided by MATLAB (functions linkage and dendrogram) was used to sort stimulus responses. To estimate the number of relevant clusters objectively, the resulting cluster tree was cut at every possible cluster number and a Hubert's Γ was calculated as

$$\Gamma = \frac{2}{O(O-1)} \sum_{i=1}^{O} \sum_{j=i+1}^{O} (S_{ij} - c) T_{ij},$$

where *O* is the size of matrix *S*, *c* is a threshold, and *T* is a binary matrix of equal size with entries

$$T_{ij} = \begin{cases} 1, \text{ if i and j are clustered together,} \\ 0, \text{ otherwise.} \end{cases}$$

The threshold was set to $c = 0.4$, as for response reliability, ensuring that only 'reliably' correlated sound responses are considered to participate in the same cluster. Again, the overall results were qualitatively similar when using a slightly different threshold, e.g. 0.3 or 0.5 (data not shown). The response modes in a given FOV were then defined as the clusters obtained for the maximal Γ-value. This clustering was highly significant ($p < 0.001$) for all neuronal populations in a FOV compared to three surrogate datasets generated by (a) shuffling the stimulus identity across trials, (b) shuffling the stimulus identity for each cell individually and (c) shuffling the cell identity for each trial (Figures S5B and S5C).

Mode-associated responsiveness of single cells

Similar to stimulus evoked sound responsiveness, we estimated whether a given cell is significantly responsive in a given response mode (see "definition of response modes"; [Figure S6A]). For each cell, we determined the rank-sum between activities from all trials associated with the mode to the same amount of spontaneous activity patterns drawn randomly from periods without sound presentation. A cell was significantly responsive in a given mode, if the p value was below 0.01 after a Benjamini-Hochberg correction against number of imaging days (four), number of modes (varying by FOV), and number of cells (21506).

Pairwise sound decoding based on full response vectors

A linear classifier (MATLAB function lassoglm with L1 regularization) was trained to discriminate between responses to two different stimuli (Figures 3D, 6D, S4D, and S10B). For the analyses where training and testing were done on the same day, cross-validation was performed by leaving out one trial. Where training and testing were done on different days, training was done with all trials of a given day and the performance of the classifier was tested on each trial of a different day. The pairwise decoding performance was then defined as the percentage of correctly classified trials, and FOV decoding performance was defined as the mean pairwise decoding performance over all pairs of stimuli.

Correlation of stimulus responses

Similarity matrices based on correlation of single trial stimulus responses (Figure S5) were first averaged for stimuli and then across FOVs. In Figure 6F the average matrix of the responses in both datasets is shown whereas in Figures S8A and S8B the matrices for the cohort of mice undergoing ACFC and the baseline cohort are shown separately.

Difference of stimulus responsiveness on day one and day seven

We computed the fraction of FOVs, in which each stimulus evoked a reliable population response (see STAR Methods section: definition of response modes) on day one and on day seven, in each of the datasets (basal and ACFC) respectively (Figure 6G). The difference of stimulus responsiveness which is plotted for each stimulus is the difference between the fraction in the ACFC cohort and the fraction in the basal condition cohort.

Difference of stimulus co-mapping to the CS + on day seven between the two experimental cohorts

Two stimuli were considered as *co-mapped*, when they evoked the same response mode (see "Definition of Response Modes") (Figure 7F). We computed the fraction of FOVs, in which each stimulus evoked the same response as the CS+, i.e. where each stimulus was co-mapped to the CS+, for both experimental cohorts separately (basal and ACFC).

Sound decoding based on a 34-fold classifier

A support vector machine was trained to discriminate between all 34 stimuli using MATLAB's built in cecoc function (Figure S4C). When training and testing was done on the same day, cross-validation was performed by leaving out one trial. When done on different days, the classifier was trained using all vectors of one day and tested with all vectors of the other day. The decoding performance was defined as the percentage of correctly classified trials.

Dependence of decoding on number of modes and stimuli per mode

We used a minimalistic model to study the dependence of sound discriminability on the number of modes and number of stimuli per mode in a FOV (Figures S10A and S10C). For simplicity, we assume all trials of a given stimulus evoke responses in the same mode (which can be the 0-mode). This model suggests that in a regime with a 0-mode as large as observed in our experimental data, an increase in both number of modes and stimuli per mode improves decoding performance.

In the model, we assume two stimuli that are mapped to the same mode are indistinguishable, but can be distinguished from stimuli mapped to a different mode (including the 0-mode). The contribution of a mode to the overall decoding performance in a FOV is proportional to its size and to a mode specific decoding factor, given by the average discriminability associated with a stimulus in this mode. This factor is large if only few stimuli are mapped to this mode, as these are distinguishable from all other stimuli. Reversely, this factor is small for modes with many stimuli, since these can only be distinguished from few other stimuli.

We cast these considerations into mathematical form to reveal how the average discriminability depends on number of modes and stimuli per mode. The decoding performance of a full FOV P_{FOV} is given by

$$P_{FOV} = \frac{1}{N_{stim}} \left(N_0 P_0 + \sum_{i=1}^{N_m} P_i L_i \right)$$

with the total number of stimuli N_{stim}, the number of stimuli evoking no response N_0, the decoding factor P_0 specific to the 0-mode, the number of different response modes N_m, the decoding factor P_i specific to response mode i and L_i, the number of stimuli mapped to response mode i. The mode specific decoding factor P_i was determined by the size of the mode L_i the stimulus is mapped to and the probabilities of its correct classification when compared to different stimuli within (P_{same}) and outside (P_{diff}) the mode:

$$P_i = \frac{(L_i - 1)P_{\text{same}} + (N_{\text{stim}} - L_i)P_{\text{diff}}}{N_{\text{stim}} - 1}$$

For simplicity, we considered the case of equally distributing the number of stimuli per mode. This results in an upper bound on the decoding performance, as can be seen by inserting P_i into the expression of P_{FOV} above. A numerical analysis revealed that this provided a reasonable approximation within the experimentally observed regime of number of modes and stimuli per mode. Thus, the expression for the decoding performance simplifies to:

$$P_{FOV} = \frac{1}{N_{\text{stim}}} (N_0 P_0 + N_m L_m P_m),$$

with the mean number of stimuli per mode L_m, the number of modes N_m. For the mode specific decoding factors, we then obtain:

$$P_m = \frac{(L_m - 1)P_{\text{same}} + (N - L_m)P_{\text{diff}}}{N_{\text{stim}} - 1}$$

and

$$P_0 = \frac{(N_{\text{stim}} - L_m N_m - 1)P_{\text{same}} + L_m N_m P_{\text{diff}}}{N_{\text{stim}} - 1}.$$

Assuming chance decoding within a given mode $(P_{\text{same}} = 0.5)$ and perfect decoding between modes $(P_{\text{Diff}} = 1)$ and a total of 34 stimuli yields Figure S10A. In general, the decoding performance increases with number of modes, while the optimal number of stimuli per mode for a given number of modes is obtained for a uniform distribution of stimuli per mode including the 0-mode, which follows from Chebyshev's sum inequality.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed using MATLAB (Mathworks, Natick, MA, USA). We used following statistical test for the given statistical analysis.

Freezing behavior

To compare freezing behavior during different time windows in a memory/generalization test session (Figures 6B, 7G, S12A, and S12B), a one-way ANOVA was performed followed by a correction for multiple comparisons with Tukey's honestly significant difference method. Stars denote significance levels: *p < 0.05.

Comparison of neuronal population dynamics between basal and fear conditioning groups

Bootstrapping was performed to test whether two mean values could origin from the same distribution (Figures 6C–6E, 7A–7D, and S10C). To do so, 10000 surrogate datasets were sampled with replacement from the original datasets. For each of the two respective values the means of the surrogate datasets were computed and compared. Next, the p value was calculated as the probability that both the surrogate from the distribution of the smaller original mean was larger than the larger original mean, and vice versa. The p values were corrected for multiple comparisons with a Bonferroni correction. Stars denote significance levels: *p < 0.05.

Difference of stimulus responsiveness and co-mapping to the CS+

Stimulus responsiveness was measured as the fraction of FOVs in which any stimulus elicited a response on day one and day seven (Figure 6G). Spearman's Rho was calculated between each pair of values for the difference in stimulus responsiveness between the two datasets and the average correlation of stimuli to the CS + on day one and on day seven. For Figure 7F, Spearman's Rho was calculated between each pair of values for the difference between the two datasets in the fraction of FOVs in which stimuli were co-mapped to the response mode of the CS+ and the average correlation of stimuli to the CS + for both datasets on day seven. The p values were analyzed for testing the hypothesis of no correlation against the alternative hypothesis of a nonzero correlation. Stars denote significance levels: *p < 0.05.

Comparison of cellular sound response profiles after sorting with data from different sessions or trials

To compare the sound evoked activity to the best stimulus assessed on a different day or a different subset of trials, a one-way ANOVA was performed followed by a correction for multiple comparisons with Tukey's honestly significant difference method. Stars denote significance levels: *p < 0.05 (Figure S3B).

Decoding comparison

To compare decoding performance of a linear classifier trained with response vectors (Figure S4D), or with maximum pooling from the best FOV of one mouse, a one-way ANOVA was performed followed by a correction for multiple comparisons with Tukey's honestly significant difference method to reject the null-hypothesis that the decoding performances of all FOVs come from the same distribution.

Comparison of mean correlations of stimulus response vectors between the experimental cohorts

To compare the mean correlations of stimulus responses in the cohort of mice undergoing ACFC and the baseline cohort (Figure S8C), we calculated their Pearson correlation coefficient and p value using a permutation test based on shuffling the order of stimuli. Stars denote significance levels: *p < 0.05.