HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2024 October 04.

Published in final edited form as:

Cell Rep. 2024 May 28; 43(5): 114172. doi:10.1016/j.celrep.2024.114172.

Auditory discrimination learning differentially modulates neural representation in auditory cortex subregions and inter-areal connectivity

Mingxuan Wang[1], **Peter Jendrichovsky**[1], **Patrick O. Kanold**[1,2,3,*]

[1]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205, USA

[2]Kavli Neuroscience Discovery Institute, Johns Hopkins University, Baltimore, MD 21205, USA

[3]Lead contact

SUMMARY

Changes in sound-evoked responses in the auditory cortex (ACtx) occur during learning, but how learning alters neural responses in different ACtx subregions and changes their interactions is unclear. To address these questions, we developed an automated training and widefield imaging system to longitudinally track the neural activity of all mouse ACtx subregions during a tone discrimination task. We find that responses in primary ACtx are highly informative of learned stimuli and behavioral outcomes throughout training. In contrast, representations of behavioral outcomes in the dorsal posterior auditory field, learned stimuli in the dorsal anterior auditory field, and inter-regional correlations between primary and higher-order areas are enhanced with training. Moreover, ACtx response changes vary between stimuli, and such differences display lag synchronization with the learning rate. These results indicate that learning alters functional connections between ACtx subregions, inducing region-specific modulations by propagating behavioral information from primary to higher-order areas.

Graphical abstract

*Correspondence: pkanold@jhmi.edu.

AUTHOR CONTRIBUTIONS

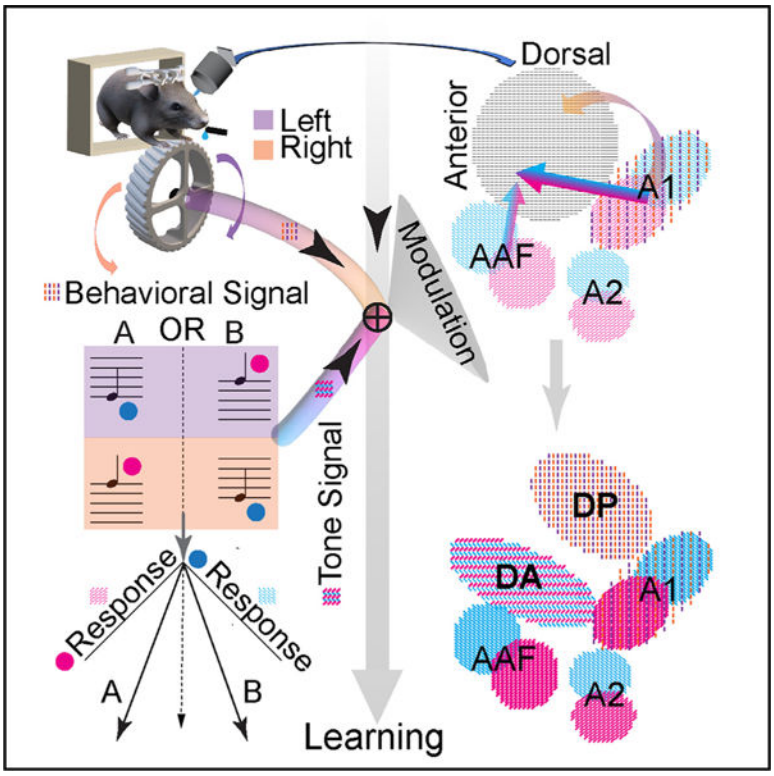
M.W. and P.O.K. conceived study. M.W. performed experiments and analyzed data. P.O.K. and P.J. supervised the project. M.W. and P.O.K. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114172.



In brief

Wang et al. develop an automated imaging system to identify mouse auditory cortex subregions that are differently recruited during discrimination learning. While stimulus- and behavioral-related signals are both present in the primary auditory cortex throughout training, higher-order regions only encode such signals post-learning. Moreover, functional connectivity between regions strengthens with learning.

INTRODUCTION

Learning and performing a sensory task, such as a discrimination task, require that sensory stimuli evoke distinct neural response patterns. This learning process can change how the brain processes sensory information.[1–8] Perceptual or associative learning of auditory tasks influences the processing and sensory representation of auditory stimuli in the auditory cortex (ACtx).[9–13] The modulation of ACtx subregions during task performance differs among the primary ACtx (A1),[14–20] the secondary ACtx (A2),[21] the anterior auditory field (AAF),[22] and other higher-order auditory regions[23–26] in multiple species such as mice, ferrets, and cats. Moreover, ACtx subregions work in concert and are thought to support hierarchical processing of sound- and task-related signals, with higher-order ACtx regions showing greater task-related modulations, similar to other sensory systems.[25,27,28] In addition, the ACtx receives motor signals[29–34]; thus, task-related motion parameters might influence ACtx responses. Hence, learning and task performance can potentially alter individual areas and their functional interactions.[35] However, most prior studies of learning-induced cortical changes have either focused only on single ACtx subregions[36,37]

or have only investigated certain time points during learning. Therefore, it is unclear how learning-induced modulations emerge in the processing hierarchy.

We hypothesized that learning to discriminate in an auditory discrimination task would enhance the differences in neuronal representation between learned stimuli. We also wanted to know when, where, and how such response saliency was improved during the learning process and speculated that the representation of behaviorally relevant signals (e.g., choice) would also change during learning.

To test our hypotheses, we longitudinally recorded ACtx activity by designing an automated head fixation training[38] and widefield Ca^{2+} imaging system that enabled voluntary imaging of mouse ACtx during the entire learning process of a 2-alternative forced-choice auditory discrimination task. We imaged all ACtx subregions and investigated the responses in each subregion and the inter-regional response correlations between subregions across different training stages and behavioral outcomes. We constructed classification models based on long short-term memory (LSTM) networks[39] to decode neural responses and identify how stimulus and behavioral choice were represented across the ACtx.

We find that across training, the tone-evoked response magnitudes increased in most ACtx subregions. Using the decoding model, we find that in the A1, throughout training, task-related stimuli and behavioral outcomes were predictable immediately from the start of the trial, and the changes across training were subtle compared with higher-order auditory regions. As the training progressed, the representation of task-related stimuli was enhanced during correct, but not wrong, trials in the anterior ACtx, and the behavioral outcomes became predictable from intermediate training stages in the posterior ACtx. Moreover, functional connectivity within higher-order auditory regions or with A1 increased. Response changes between the two learned stimuli were asymmetric, and motion signals in the ACtx contributed to these asymmetrical response changes. In addition, we find that such neural response changes happened after the behavioral performance changes and that the response asymmetricity displayed lag synchronization with the learning rate. Altogether, our results suggest that learning a forced-choice auditory discrimination task induces region-specific modulations of both tone- and behavioral-related responses in the higher-order ACtx and alters functional connections between ACtx subregions, presumably by the task-related movement signals.

RESULTS

Mice learned a forced-choice auditory discrimination task during automatic head fixation and longitudinal widefield imaging

To investigate the changes in ACtx responses during learning, we constructed an automated head fixation training system[38] combined with a widefield Ca^{2+} imaging system (Figure 1A, top). This system allows longitudinal widefield Ca^{2+} imaging across the ACtx (>3 × 3 mm^2). The advantage of automated head fixation training is that animals are free to engage in the task, and it also reduces the potential influence of the experimenter.[40,41] To ensure consistent imaging across sessions and days, the focal position for widefield imaging was automatically determined via an integrated autofocus system before the start of each training

session for each mouse, and the imaging fields were matched (see STAR Methods and Figure S1) after imaging. We used mice expressing the genetically encoded Ca^{2+} sensor GCaMP6s[42] across all layers of the ACtx,[43] including layers 2/3 and 4 (Figure S2), and analyzed fluorescence responses during voluntary task performance.

We trained mice ($N = 8$) on a 2-alternative forced-choice tone discrimination task (see STAR Methods) in which they had to turn a wheel to the left (positive degrees) when a low-frequency (10 kHz) tone was presented and to the right (negative degrees) when a high-frequency (40 kHz) tone was presented (Figure 1A, bottom). We also trained a separate cohort of animals ($N = 6$) with the opposite association (i.e., turn the wheel to the right when hearing a low-frequency tone, and vice versa) to exclude the potential effects of the asymmetric stimulus-motor association.

The ACtx in trained animals encodes both stimulus and behavioral choice,[14,15,44] and disentangling the encoding of each and potential interactions between ACtx subregions requires behavioral paradigms that minimize the likelihood of mixing stimulus-related signals with choice- or movement-related signals.[14,15,44,45] Thus, a hold period (1/3 s) was present in each trial, during which mice were trained not to execute any movement. Although the hold period was not long enough to cover the whole duration of the tone presentation, this hold period still enabled us to partially separate tone-related and movement-related signals. The period right after the hold phase but before the tone offset was considered the move phase. The training period lasted 25 days, and in most of our later analyses, we divided the whole training period into five stages of 5 days each to reduce potential variabilities across training days and to ensure enough trials (~500) in each stage for model training.

Rotary trajectories of an example mouse are shown for the first and last training stages (Figure 1B). Turning directions of the wheel were distributed randomly across two frequencies at the first training stage but were well separated at the last training stage, in that rotary trajectories went up for most low-frequency trials and down for high-frequency trials.

To determine whether mice had learned the task after training, we calculated the sensitivity index (see STAR Methods), d', which measures discriminability,[46] for each mouse on each training day (Figure 1C). A threshold of 1 was chosen for d' to indicate the ability of mice to discriminate between low- and high-frequency tones. On the last training day, all mice showed a higher likelihood of giving correct over wrong responses, with a discrimination ability of 1.64 ± 0.53 ($N = 8$, mean ± standard deviation) (Figure 1C). The cohort trained with the opposite association achieved a similar performance (Figure S3A). These behavioral results indicate that all mice could discriminate between low- and high-frequency tones after training, with high response rates, and most correct decisions were made close to the start of the choice window (Figure 1D). The number of trials performed by each mouse was also kept consistent over training (Figure 1E).

The tone response magnitude in multiple ACtx regions changed across training

We first sought to identify which ACtx subregions showed altered responses over training. We identified ACtx subareas[47,48] guided by the tone-evoked cortical responses on the first training day (Figure 2A). We used miss trials to avoid movement- and decision-related confounds (see STAR Methods and Figure S4). We then derived pixelwise average fluorescence (ΔF/F) maps (see STAR Methods) for low-frequency correct (LC) or high-frequency correct (HC) trials at the first and last training stages during the entirety of the hold or move phase (Figure 2B). The neural responses during the hold and move phases changed in most ACtx subregions from the first to the last training stage, especially in AAF, the non-response DA, and the dorsal posterior auditory field (DP) (Figure 2B, yellow arrows). Thus, training changed the tone-evoked responses in multiple ACtx subregions, especially in the dorsal regions.

The changes in ΔF/F maps (Figure 2B) suggest that the magnitude of responses during different phases might have changed at various training stages. Thus, to investigate the responses in more detail, we examined the average fluorescence (ΔF/F) traces of LC or HC trials for each ACtx subregion at each training stage (Figure 2C). At the second training stage, the A1, A2, and AAF had decreased response magnitudes during both the hold and move phases in LC trials compared to the first training stage. In contrast, we did not observe such decreases for HC trials. Meanwhile, in HC trials, most ACtx subregions, especially the A1, AAF, and DA, displayed increased response magnitudes during the move phase at later training stages. We observed similar response increases for the cohort with the opposite training association except that responses in LC trials increased more compared to HC trials (Figure S3B). These results suggest that learning our auditory task modulates tone-evoked response magnitudes in most auditory subregions, consistent with prior studies.[14,24–26] In addition, our results indicate that asymmetrical changes exist in the response magnitudes to low and high frequencies.

Representation of the task-related tone frequencies was enhanced in most ACtx subregions after training

Animal performance improved over training, and we speculated that the altered responses in the ACtx with training reflect the emergence of a more discriminable representation of the learned stimuli. To examine this notion and if such representation varies across subareas, we built a classifier based on LSTM networks (Figure 3A) to decode cortical responses and predict the presented tone frequencies (see STAR Methods), the performances of which were used to interpret the representation of the cortical responses to the learned stimuli. We chose bidirectional LSTM architecture due to its strength in handling temporal data and learning long time dependencies.[49]

We first focused on LC and HC trials from each training stage for the model training and testing. Pixels of each ACtx subregion were considered features of their corresponding subregion, and the fluorescence values of each pixel during specified time intervals were treated as time series data. For the training of models for each animal (8 mice in total), time series data of 5 consecutive frames (1/6 s in each model, from −1/2 to 1 s with regard to [w.r.t.] the tone onset, 9 time intervals in total) from each ACtx subregion (14

ACtx subregions as well as the whole ACtx to comprise 15 regions in total) and from each behavioral training stage (5 training stages) were used in a 10-fold leave-one-out cross-validation process (10 times training and testing on non-overlapping testing datasets). We used the area under the receiver operating characteristic curve[50] (AUC) as the evaluation metric. We averaged AUCs across different validation processes and time intervals for each mouse and each subregion and at each behavioral training stage. AUC is an appropriate evaluation metric because it balances the trade-off between sensitivity and specificity at the best-chosen threshold, making it exclusively suitable for evaluating model performance on an imbalanced dataset.[51]

As the training progressed, the model performance showed an overall increase in the prediction accuracies for the presented tone frequencies (Figure 3B; Table S1). Similar results were also found for the cohort with the opposite training association (Figure S3C). Responses from the A1 showed higher decoding performance for the presented tone frequencies at the first training stage and displayed consistent prediction performance across the training period. In contrast, during the first training stage, responses from the higher-order ACtx regions (e.g., DA and the non-response DP [DP_n]) did not yield performance results as good as the A1. However, the model performance results increased over the training and peaked at the last training stage, especially in DA. The model performance results mirrored the changes in the response magnitudes across the behavioral training stages. Thus, learning an auditory discrimination task increases the representation of the presented tone frequencies in the neural activity of the higher-order ACtx areas.

Wrong trials showed a less faithful representation of the task-related tone frequencies than correct trials

To investigate if the representation of the learned stimuli varies between different behavioral outcomes, we trained the models similarly on low-frequency wrong and high-frequency wrong trials. Because the numbers of correct and wrong trials differed across training, we matched the training sample size between correct and wrong trials for each model to eliminate the effects of sample sizes on the performance of the models and retrained the models on LC and HC trials with the matched data size for each class. We chose 5-fold leave-one-out cross-validation here to ensure that data points from all classes were presented in the testing dataset during each validation process and that AUCs could be calculated as evaluation metrics for model performance.

We find that the differences between the performance of models trained on correct and wrong trials differed across different behavioral phases at later training stages, especially in the anterior part of the ACtx, as suggested by the pairwise comparison (Figure 3C; Table S2). Similar results were also found for the cohort with the opposite training association (Figure S3D). Our results indicate that the representation of the task-related tone frequencies is unreliable in wrong trials at later training stages, and we speculate that this unreliability (i.e., lack of or incorrect information about the stimulus) leads to the wrongly performed trials.

A1 activity predicted behavioral outcomes across all training stages, while DP_n activity predicted behavioral outcomes at later stages

A1 neurons can encode behavioral outcomes.[14] To identify if other ACtx subregions encode behavioral outcomes and when such information might emerge, we next aimed to decode behavioral outcomes from ACtx responses. We combined both correct and wrong trials during model training. Models were trained to predict behavioral outcomes of either correct or wrong trials separately in each animal (8 mice in total), for each ACtx subregion (14 subregions as well as the whole ACtx for a total of 15 regions), from each behavioral training stage (5 training stages), and from time series data of 10 consecutive frames (1/3 s in each model, from −1/3 to 1 s w.r.t. the tone onset, 4 time intervals in total). We also chose 5-fold leave-one-out cross-validation here for the same reasons as described above. Average AUCs in predicting behavioral outcomes (Figure 3D) were plotted for each mouse, each subregion, and each time interval of the data and at each behavioral training stage. We compared the performance of models trained on each time interval within the 0 to 1 s period with those trained on the −1/3 to 0 s period to determine if any model could achieve a performance accuracy higher than the chance level (Table S3).

The A1 showed a consistent and reliable decoding capability of behavioral outcomes across all behavioral training stages (Figure 3D). Meanwhile, as the training progressed, DP_n displayed an increasing predictability of behavioral outcomes (Figure 3D; Table S3). Similar results were also found for the cohort with the opposite training association (Figure S3E). We then calculated the differences between the performance of models trained on data from each ACtx subregion and the performance of models trained on all ACtx data to compare the predictability of behavioral outcomes among different ACtx subregions. We observed significant differences between the performance of models trained on data from the DA_n and those trained on all ACtx data across all time bins at the last training stage (Table S4), which suggests that the DA_n does not present as much information about behavioral outcomes as other subregions.

Altogether, these results suggest that A1 reliably encodes information about behavioral outcomes throughout the training stages, and such information is encoded in the DP_n at later training stages.

Enhancement of inter-regional correlations of the tone-evoked responses with higher-order areas

Auditory areas do not work in isolation, and we speculated that in addition to the changes in specific regions, task learning might also change the interaction between ACtx subareas. We thus calculated the activity correlations between ACtx subareas during the hold or the move phase.

At the first training stage, the A1, A2, and AAF are highly correlated in both correct and wrong trials (Figure 4A), especially during the hold phase. In contrast, the correlations of the dorsal auditory regions between each other and with other areas were relatively low. As the training progressed, correlations between the primary and the higher-order areas increased after slight decreases in the second training stage (Figure 4A) in correct trials, especially

between the A1 and DA, the AAF and DA, and the A1 and DA_n during the hold phase and between the DA and DP_n during the move phase (Figure 4B, left and right; Table S5). Meanwhile, in wrong trials, despite similar increases in the correlations between the A1 and DA and the AAF and DA, we also found correlations between the A1 and AAF and the DA and DP, as well as the A1 and DP_n, to be increased in the hold phase and the AAF and DA in the move phase (Figure 4B, middle and right; Table S5). Results were similar for the cohort with the opposite training association (Figure S3F).

These results indicate that during correct trials, tonal-responsive subregions decorrelate with higher-order subregions during early training stages, consistent with the observed decrease in the response magnitudes in the A1 but not in the dorsal medial auditory field (DM), the DA, or the DP (Figure 2C). At later training stages, the higher-order subregions showed increased response magnitude (Figure 2C) and correlated activity with the A1 in both correct and wrong trials. Thus, our results indicate that training has differential effects on network processing in the ACtx, and the primary and higher-order regions become more functionally connected during learning.

Behavioral-related motion signals attributed to the asymmetries in the response changes

So far, our results have shown that learning affects both subregional responses and inter-regional correlations, and we noticed asymmetries in the response magnitude changes between the two stimuli. We then asked where those asymmetries arise from and whether such asymmetries are related to the learning process. Given that our task is asymmetric, which requires the mice to spin the wheel in a specific direction, behavioral bias, e.g., ‘‘handedness,’’[52] could contribute to our results and introduce asymmetries in neural response. We therefore first investigated whether the differences in the response changes between the two stimuli could be partially due to a behavioral bias of the animals. Hence, for each animal, we quantified its behavioral bias on each training day by calculating the difference between the hit rates for low- and high-frequency trials, and we quantified the response asymmetry on each training day by calculating the average ΔF/F between LC and HC trials for each ACtx subregion during either the hold or the move phase. Across all mice on all training days, we found no strong correlations between the behavioral bias and the asymmetries in response changes between the two frequencies (Figure S5). These results indicate that the observed response bias is not due to the animals’ intrinsic behavioral bias.

Given that movement signals are present in the ACtx,[29–34] and since our task had an asymmetric association between tone frequency and spin direction, we speculated that such signals could contribute to the asymmetrical response changes we observed, as the control of sensory processing was found to be shaped by the task reward structure in addition to the required sensory discrimination as a neural coding strategy to maximize the discriminability between two learned stimuli.[17] For example, the stereotyping of behavioral movements could induce response differences between the two sound frequencies. Thus, we investigated whether the asymmetrical changes always happened to one stimulus (e.g., only increased responses to high-frequency tones) or whether such changes were driven by a specific behavioral motion (e.g., only increased responses to the right-wheel-turning trials).

To test this possibility, we trained a separate cohort of animals ($N = 6$) with the opposite stimulus-wheel-spin association (i.e., turn the wheel to the right when hearing the low-frequency tone, and vice versa). From the average ΔF/F traces of LC or HC trials in each ACtx subregion across training, we notice that, contrary to the previous results, tone-evoked response magnitudes increased more in LC trials than in HC trials (Figure S3B). To measure the asymmetricity of the response changes in either case (stimulus-wheel-spin association), we first calculated the averaged ΔF/F traces from either LC trials or HC trials and averaged them during each time interval (the hold or the move phase) for each animal in each ACtx subregion and at each training stage. We visualized the differences in responses to low and high frequencies by plotting the average response in LC versus the average response in HC (Figures 5A and 5B, each dot corresponds to data from one animal at each training stage). For each training stage, we then averaged the dots across animals (Figures 5A and 5B, larger dots) and performed linear regression across stages 1–5. In the regression plot, we pointed the direction from the dot corresponding to the first training stage toward the dot for the last training stage and regarded the vector direction w.r.t. the line y = x as the asymmetricity of the response changes (Figures 5A and 5B, vectors are shown by black arrows, and line y = x is shown in dashed line). We find that the asymmetries in the tone-evoked responses were reversed when the animals were trained on the opposite association (Figure 5B). Thus, the wheel-turning direction matters, and we conclude that motion signals in the ACtx contribute to the asymmetrical response changes between learned stimuli that we observed in the ACtx.

The response discrepancy between different stimuli correlated with the learning rate

Given that the task-related movements existed throughout the training but the motion-affected response changes happened only at specific training stages (Figures 2, 3, and 4), we asked whether learning is the driving force for the tone-evoked responses to be modified by the motion signals to improve the discriminability between two learned stimuli. To quantify the response discrepancy for two stimuli (see STAR Methods), we calculated the average absolute difference between the mean ΔF/F traces of LC and HC trials during each time interval (the hold or the move phase) for each animal in each ACtx subregion and on each training day (Figure 6A, left, green lines).

If this response discrepancy were necessary for good performance or would have been induced by learning, we would expect it to correlate with the performance or the learning rate over training. Therefore, we defined the learning rate by the changes in behavioral discriminability ($\Delta d'$) and plotted it together with the response discrepancy (Figure 6A, magenta lines). We find that the changes in the response discrepancy displayed high synchronicity with the learning rate, with a lag of more than 3 days (Figure 6A, right). Similar results were also found for the cohort with the opposite training association (Figure S3G). To evaluate this apparent relationship, we then calculated the cross-correlation between the response discrepancy and the learning rate (Figure 6B). We find that response discrepancy during the hold phase showed a higher lagged correlation with the learning rate in the A2, DM, and DP with a lag of around 5 days, while during the move phase, a higher lagged correlation was observed in the DM and DA_n with a shorter lag of approximately 3 days.

Altogether, our results suggest that in the ACtx, instead of the changes in neural response amplitude leading to increased behavioral performance, the neural response changes lag the behavior changes.

ACtx showed persistent learning-induced changes not due to movements

We find that motion signals in the ACtx contributed to the asymmetrical response changes between learned stimuli that happened after the behavior changes (Figures 5 and 6) and that the representation of task-related stimuli was enhanced after training (Figure 3). Given that the movements became stereotyped (Figure 1B) as the animals learned the task, it is possible that such motion signals could be the main cause of the changes we observed (e.g., the enhanced neural representation of the stimuli could be just due to the distinct motion signals in the ACtx for different movements instead of the solely tone-related responses being differentiated for different stimuli). We wanted to know whether the lagged response changes reflected any long-term alternations to the tone-evoked responses in the ACtx that were irrelevant to the behavioral movement.

We thus investigated the average tone-evoked responses from miss trials (Figure 7A) and found that although the responses were weaker for miss trials compared to correct trials, the same trend of response amplification after training was still noticeable. Like the original paradigm (Figure 2C), responses to high-frequency tones in both the A1 and DA_n increased after training in miss trials. Thus, the increase in sound-evoked responses over training existed even without the presence of movements (response changes to the passive hearing of the tones). We next investigated the changes to the inter-regional response correlation in miss trials (Figure 7B). We find that, like the changes in correct trials, the primary and higher-order regions also became more functionally connected in trials without movements. We then tested if the changes in the neural representation of the stimuli were dependent on the movements. We again used the decoding models trained from correct trials to decode tone frequencies from miss trials (Figure 7C). At the last training stage, subregions that were found to have increased representation of tone frequencies (e.g., DA_n) in correct trials also showed such increases in miss trials, and the performances in correct and miss trials were similar (Table S6). Our results suggest that the learning permanently alters the response and functional connectivity of the ACtx subregions and results in a more discriminable representation of the learned stimuli and that such changes are not merely due to the encoding of motion signals.

In conclusion, these results suggest that during the learning of an auditory discrimination task, behavioral-related motion signals emerge in the ACtx. Correlations between hearing a specific tone and turning the wheel in a specific direction are gradually established (or stereotyped). Such modulation eventually alters (enhances) the ACtx circuits to better represent (or perceive) the tone frequency.

DISCUSSION

We investigated learning-induced changes in ACtx processing. We find that learning a forced-choice auditory discrimination task recruits the higher-order ACtx, induces region-

specific modulations of both stimulus- and behavioral-related responses in the ACtx, and increases functional connections between primary and higher-order areas.

We here used an automatic head fixation[38] and imaging system. Automation removes variability introduced by the experimenter and lowers stress in the experimental subject.[40,41,53,54] By restricting the total amount of daily training time per mouse, our training paradigm allows the efficient training of cohorts of mice. Our imaging system enables us to record the neural activity of the whole ACtx in one (left) hemisphere. In contrast to previous work focusing on single ACtx subareas,[14–20,22,25,36] the advantage of mesoscale imaging is that we could track learning-induced inter-regional changes across the ACtx longitudinally. By observing all subregions simultaneously we revealed the change in individual subregions and the correlated activity between subregions.

We identify training-related changes in the primary and higher-order ACtx. We find that different ACtx subregions displayed diverse characteristics in the change of response profiles in behavioral contexts while learning an auditory discrimination task. Using LSTM models, we show that the representation of stimuli involved in the training and the behavioral outcomes varied among different ACtx subregions, indicating that learning an auditory decision-making task involves changes in multiple ACtx areas and that information about the stimulus and the behavioral outcome is represented differently across ACtx subregions.

In addition, we find that activity correlations between the primary and higher-order areas increased over training. Given that correlated activity can indicate functional connections, our finding suggests an increase in the inter-areal connectivity patterns and that training can alter how sounds are processed across the ACtx by changing ACtx circuits. This circuitry change was further supported by our results showing the enhanced representation of the learned stimuli in miss trials, which suggests long-term changes to the intrinsic sound processing circuits in the ACtx.

Moreover, our results show that the tone-evoked response magnitudes changed asymmetrically for the two learned stimuli in most ACtx subregions and that the associated wheel-turning direction determined the asymmetry direction. Consistent with other studies,[29–34] our results confirm the existence of behavioral-related motion signals in the ACtx. However, as one would expect, such motion signals should occur concurrently with solely tone-related signals as the animals were performing the task, and if the observed asymmetrical response changes were merely due to the integration of the motion signals, it should happen synchronically and uniformly as long as the animals were performing the correct trials. Instead, we discover that the changes in the response discrepancy (asymmetricity) between two learned stimuli lagged the changes in the learning rate for 3 to 5 days. Since the ACtx is actively involved in the consolidation of long-term memory, as supported by molecular studies,[55,56] we speculate that once the animals start to ‘‘know’’ the task (acquire the strategy), a learning signal might be present together with the signals from the turning of the wheel in a specific direction and the solely tone-evoked responses in the ACtx that result in the lagged neural response changes.

In terms of the functional specialization of ACtx subregions, we first notice that stimuli and behavioral choice could be better predicted from the A1 than higher-order areas from the onset of training. Moreover, in the A1, we find that the changes of decreasing response magnitudes happened at early training stages. Hence, they could represent neuronal adaptations to the repeated stimuli.[36,57–59] Behavioral outcomes can also be predicted from the A1 at all training stages, consistent with prior reports.[1,14,60] While it has been thought that there is hierarchical processing of stimulus- and task-related signals across the ACtx with higher-order areas showing larger task-related influences,[61–64] our results instead support a more parallel processing view, suggesting that the A1 serves as the primary site for early auditory processing that is sensitive to new auditory experiences and behavioral signals might be propagated from the A1 to higher-order areas to facilitate the performing of the task, e.g., to improve the discriminability of the tones or to store memory,[65–68] as the animals become more proficient at the task.

In the posterior part of the ACtx, e.g., the DP, we observe high predictability of the behavioral outcomes started from intermediate training stages, consistent with reports that the DP is involved in auditory perceptions in mice[69] and humans.[70,71] In the anterior part of the ACtx, e.g., the DA_n, we find that the asymmetrical response changes were highly correlated with the behavioral learning rate, and the representation of tone-related frequencies was improved after training. The origins of the tone-frequency signals in those anterior parts might be the A1, as suggested by the increased functional connectivity. We further speculate that this enhanced representation of learned stimuli is achieved by the asymmetrical modulations of the tone-evoked responses by the behavioral-related motion signals, inferring the potential role of those anterior subregions in the ACtx in the refining of auditory perception through the integration of sensory and motor signals initiated by learning.

Limitations of the study

Our behavioral paradigm only partially separated solely tone-elicited signals (the first 1/3 s) and motion-related signals (the remaining 2/3 s of the trial contain both tone-elicited and motion-related signals). A more segregated behavioral paradigm for the hold and move phases might result in an even better separation of tone-related and motion-related signals. However, it is challenging to train the mice to withhold movements for a prolonged period without any external cues (e.g., visual cues would introduce confounding factors). We used widefield imaging to record calcium signals from Thy1– GCaMP6s mice. Due to the nature of one-photon imaging, the overall signals likely reflect a mixture of neural responses from multiple cortical layers. The density of labeled layer 2/3 cell bodies is high in our Thy1 line (Figure S2). Thus, these cells likely contribute a large fraction of the signals we recorded. However, apical dendrites of deeper neurons, e.g., layer 5/6, reach into superficial layer 1 and could also contribute to the imaged signals. Using layer-specific calcium indicator expression could aid in understanding better the laminar origins of the learning-induced changes in the future. In addition, auditory processing involves two hemispheres, and training-related changes might show hemispheric differences,[72] which we did not investigate here. In our functional connectivity analysis, we only correlated ACtx responses in different subregions for different stimuli or behavioral outcomes but did not take other factors, such

as arousal, motivation, working memory, attention, etc., into account. Lastly, as correlation does not imply causality, loss-of-function (e.g., optogenetic inactivation) or stimulation experiments could be performed to confirm the role of the ACtx subregions that were found to have learning-induced changes in our current study.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Patrick O. Kanold (pkanold@jhmi.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability

- Calcium imaging data and behavioral data have been deposited at the JHU Research Data Repository and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- All original code has been deposited at the JHU Research Data Repository and is publicly available as of the date of publication. The DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT PARTICIPANT DETAILS

Animals—For calcium imaging and behavioral training, we used 7 male and 7 female adult mice that were the F1 generation of C57BL/6J-Tg(Thy1-GCaMP6s)GP4.3Dkim/J (JAX# 024275) crossed with B6.CAST-Cdh23Ahl+/Kjn mice (JAX#002756) in this study. The genotype of the mice we used is Tg(Thy1-GCaMP6s) positive and heterozygous for $Cdh23^{Ahl+}$. The F1 generation has minimal hearing loss throughout their lifespan.[73] These 14 mice came from 6 litters. Littermates of the same sex were randomly assigned to experimental groups, and we balanced the number of male and female mice in each group to avoid any potential influence of sex. The mice used in this study ranged from 2 to 3 months old when training started. All mice were housed in a 12-h reverse light/dark cycle room in the institutional animal colony before and after the experiments, and all experiments were conducted during their dark cycle. For the immunohistochemistry experiment, we used a 5-month-old female mouse with the same strain and genotype, just as the mice we used for the calcium imaging and behavioral training experiments. All protocols and procedures were approved by the Johns Hopkins Institutional Care and Use Committee.

METHOD DETAILS

Surgery—All surgeries were performed under anesthesia using isoflurane (4% induction, 1.5% maintenance, VetOne). To prevent brain swelling during cranial window implant, 0.1 mL of dexamethasone (2 mg/mL, VetOne) was injected subcutaneously 0.5 to 1 h before surgery.[74] Right before each surgery, 0.1 mL of dexamethasone (2 mg/mL) and 0.05 mL of

atropine (0.4 mg/mL) were injected subcutaneously. Hair on the top of the head was first removed, followed by removing skin and soft tissues underneath using disinfected scissors and scalpel blades. The dorsal and caudal sides of the skull were then exposed by removing muscle on the surface. Cranial window surgery was then performed on the left side of the skull. The center of the window location was determined by landmarks on the skull. A circular area of the skull with a diameter approximately equaling 3.5 mm was then removed around the center by a dental drill. After cleaning the exposed brain surface using sterile saline, a custom-made cranial window was placed on top. The window was made with a layer of 3 mm round coverslips (catalog #64–0720-CS-3R, Warner Instruments) stacked at the center of a 4 mm round coverslip (catalog #64–0724-CS4R, Warner Instruments) and secured with optic glue (catalog #NOA71, Norland Products). The edge of the window was then sealed with Kwik-sil (World Precision Instruments), and dental cement (C&B Metabond) was then applied to secure the window to the skull.[74] To prepare mice for automated head-fixation training and imaging, custom-designed headplates were mounted to their skull[38] after the cranial window surgery. Repeated measurements of the z axis reading of the lambda and bregma landmark on the skull were then made to adjust the horizontality of the head until the difference between readings of the two landmarks was less than 200 μm. A headplate was then attached to the skull by super glue such that both the midline of the headplate and the midline of the head were aligned, and the front of the headplate was aligned with the lambdoid sutures.[38] Dental cement was used to cover the exposed area of the skull (C&B Metabond). To implant the RFID tag, a small incision was made over the skin covering the right abdomen, and the RFID tag was inserted before suturing the skin. After the surgery, 0.05 mL cefazolin (1 g/vial, West Ward Pharmaceuticals) and 0.01 mL carprofen (0.5 mg/mL, Zoetis) per gram of mouse body weight were injected subcutaneously, and the mouse was placed under a heat lamp for recovery of 30 min before being placed back in the home-cage. The same amounts of cefazolin and carprofen were also injected subcutaneously 3 days, 7 days, and 14 days after surgery. Training began 14 days after surgery was performed.

Automated training with one-photon imaging—The automated training apparatus was assembled in the same way as in our previous study.[38] Mice were put into the training cage (Figure 1A, top) and were allowed to voluntarily get head-fixed for training. Once a mouse got into the tunnel, the unique RFID of the mouse could be read by an RFID scanner to recognize the identity of each mouse during training. Automated head-fixation would then be initiated after the presence of the mouse head was detected by beam break sensors (Adafruit, product ID 2168).

A CMOS camera (Thorlabs CS505MU) was used to record fluorescence signals at 30 Hz, 16-bit 488 × 408 pixel resolution, with the excitation by blue LED (Thorlabs M470L5) driven by T-Cube LED Driver (Thorlabs LEDD1A). A Plano-Convex lens (Thorlabs LA1951-A) was placed between the LED and an excitation filter (Chroma ET470/40) where excitation lights converged and were later reflected to the tissue by a dichroic mirror (Thorlabs MD499) placed in an Epi-Illuminator module (Thorlabs WFA2002) followed by an objective (Olympus UPlanFl 4×/0.13). The emission light is passed through the objective, the dichroic filter, an emission filter (Chroma AT535/40m), and a camera tube

(Thorlabs WFA4102–0.5×) before entering the CMOS camera. The whole imaging parts were mounted at 45° on a motorized breadboard driven by a T-Cube DC Servo Motor Controller (Thorlabs TDC001) with an Actuator (Thorlabs Z812B).

Matching of the region of interest (ROI)—Determination of imaging plane: The focal position for widefield imaging was automatically determined after automated head-fixation was done. For a newly registered mouse, the imaging setup is first moved for a range of 10 mm with a 1 mm step size, and a single widefield image will be recorded at each position (Figure S1A, bottom, five example images taken at each position as indicated by orange arrow). An index that measures the degree of focus was then calculated for each image (Figure S1A, top, plot showing index calculated at each position). Widefield images were normalized individually by dividing the subtraction of the minimum intensity within an image from each pixel by the difference between the maximum and minimum pixel intensities. Histogram equalization was then performed for each image to enhance the contrast. The concentric area with 30% of the width and 30% of the length of the image was then cropped as a $m \times n$ matrix X for calculating the index. Fast Fourier Transform was then performed:

$$Y_{p+1,q+1} = \sum_{j=0}^{m-1} \sum_{k=0}^{n-1} e^{-2\pi jpi/m} e^{-2\pi kqi/n} X_{j+1,k+1}$$

where i is the imaginary unit and p, q are indices running from 0 to $m - 1$ and $n - 1$; respectively. The Fourier transform Y was then rearranged to Y' by swapping its first and second quadrants with the third and fourth accordingly to shift the zero-frequency component to the center. The index was eventually computed as:

$$\text{index} = \sum_{j=\lfloor m/2-r \rfloor}^{\lfloor m/2+r \rfloor} \sum_{k=\lfloor n/2-r \rfloor}^{\lfloor n/2+r \rfloor} |Y'_{j,k}|/(2r+1)^2$$

where r controls the sensitivity of the index to the degree of focus and was set to 8 during application to achieve the best performance. The step size was then changed to 0.05 mm to search for the position with the highest index within the range of – 0.5 mm to +0.5 mm with respect to the position with the highest index in the previous search. The position with the highest index was now considered the focal position for widefield imaging of the mouse (Figure S1A, indicated by the middle vertical arrow), and was recorded and updated before each later imaging session as the center position for 0.05 mm step size focal position search. The focusing process takes less than 10 s once the initial searching position is given.

Matching of the imaging field of view: Unlike focal position (z axis), which was determined before each imaging session, the field of view (x axis and y axis) of each imaging session was separately matched afterward. For a typical imaging session, motion correction[75] was first performed, and an average image was computed (Figure S1B, upper left, first row). The same normalization and histogram equalization as discussed in the previous paragraph was then performed on the average image (Figure S1B, upper left, second row), followed

by convolution with a Gaussian filter with a standard deviation of 2 (Figure S1B, upper left, third row) to eliminate effects of minor vasculature changes over time to the matching process while only focus on major vasculatures as landmarks for the field of view matching to the template. The template image underwent the same pre-processing to yield the template $m \times n$ matrix, T. The pre-processed average $m \times n$ image, X, was then moved around the center of the template matrix (Figure S1B, lower left) with shifts in x axis, ix, and y axis, iy, both ranging from −30 to 30 pixels in the application. Shifts with the minimum average absolute intensity differences of the overlayed area between X and T (Figure S1B, upper right, dark blue point in the heatmap) were used to translate the motion-corrected imaging data $I(x, y, t)$ to match the field of view of the template (Figure S1B, lower right):

$$\left\{ I' = I(x + \Delta x, y + \Delta y, t) : \{\Delta x, \Delta y\} = \underset{ix, iy}{\operatorname{argmin}} \sum_{j = \max(1, 1 + iy)}^{\min(m, m + iy)} \sum_{k = \max(1, 1 + ix)}^{\min(n, n + ix)} \left| T_{j,k} - X_{j - iy, k - ix} \right| / \left[(m - |iy|) \cdot (n - |ix|) \right] \right\}$$

Behavioral training paradigms—Mice were water-restricted throughout the training. To ensure the survival and normality of the mice, we would deliver them free water if the body weights were less than 80% of their initial weights at the end of each training day. Mice were first put into the training chamber for 3 to 5 days, with no actual training taking place for acclimation. During the first 1 to 2 days, mice were rewarded with water at the end of the tunnel without head-fixation. For later 1 to 3 days, mice were head-fixed for every entry, and 5 drops of water (opening the pump for 200ms for 1 drop, approximately 20 μL per water drop) were given with 2 s delay between each drop. On each acclimation training day, each mouse would stay in the chamber until it consumed 1 mL of water, which was determined by the change in its body weight. Once a mouse got used to the training chamber and head-fixation (was able to consume 1 mL of water within an hour), auditory discrimination training would begin.

For the tone discrimination task, mice were trained to turn the wheel to one (e.g., left) direction when hearing a 70 dB sound pressure level (SPL) 1 s 10 Hz fully amplitude-modulated low-frequency (10k Hz) tone and to the other direction (e.g., right) for 70 dB SPL 1 s 10 Hz fully amplitude-modulated high-frequency (40k Hz) tone. All sound waveforms were generated by NI-USB6215 (National instrument) and routed to an ED1 speaker driver (Tucker-Davis Technologies) for presentation using an ES1 speaker (Tucker-Davis Technologies). Tones of different frequencies were calibrated *in situ* with a B&K microphone (Bruel & Kjaer 4944-A). Choices were considered made once the wheel was turned to pass a threshold of 10° in either direction. Any decisions made during the hold period (0–1/3 s from tone onset) were considered early response trials where a buzzer was activated for 1 s as punishment. Correct trials were rewarded with a drop of water (opening the pump for 200ms, approximately 20 μL) right after tone offset or when a decision was made, whichever came later, whereas the wrong trial was punished with a prolonged time-out duration from 4 s to 20 s. Trials with no decision made during the choice period (4 s after the end of the hold period) were considered miss trials.

A training session was delivered once the automated head-fixation and focusing process were finished, and imaging triggers were synchronized with behavioral triggers. The duration of the training sessions consistently changed during the training process. Initial minimum and maximum session durations were 2 min. The minimum session duration would increase with a step size of 0.25 min for every 5 instances of head-fixation (training sessions) until it reached 3 min, and the maximum session duration would increase with a step size of 0.5 min for every 5 instances of head-fixation (training sessions) until it reached 24 min. A training session would end only if it had passed the maximum session duration or 5 consecutive unrewarded trials (early trials, wrong trials, and miss trials) were presented with the session time exceeding the minimum session duration. Within a training day, each mouse was put into the training chamber for approximately an hour, during which the mouse was allowed to freely initiate multiple training sessions until the allocated training time ended. During each training session, low-frequency trials and high-frequency trials were presented randomly with a minimum base inter-trial-interval (ITI) of 4 s plus time for reward consumption (2 s) or prolonged time-out duration for wrong trials (from 4 to 20 s) if applicable. A new trial would only be delivered when it passed ITI from the end of its previous trial, and the animal held still for at least 0.5 s (changes in rotary readouts not passing a threshold of 5° in either direction within 0.5 s).

Immunohistochemistry and imaging—An adult (5-month-old) mouse (F1 generation of C57BL/6J-Tg(Thy1-GCaMP6s)GP4.3Dkim/J (JAX# 024275) crossed with B6.CAST-Cdh23Ahl+/Kjn mice (JAX#002756)), just as what we used in our study, was transcardially perfused with cold phosphate buffered saline (PBS) and 4% paraformaldehyde (Electron Microscopy Sciences) under deep anesthesia (Fluriso, VetOne). The perfused brain was fixed in 4% paraformaldehyde for 24 h before being transferred to a PBS solution. The brain was sectioned with a vibrating microtome in the coronal plane at a thickness of 50 μm and stored in PBS until use. To identify GFP-immunopositive cells, we performed immunohistochemistry using standard protocols. A blocking solution was prepared with 5% normal goat serum (Cat no. 5560–0007, SeraCare) and 0.3% Triton X-100 (Sigma). We incubated selected brain slices in the blocking solution at room temperature for 90 min before incubating overnight in chicken-*anti*-GFP (1:1000, ab13970, Abcam) at 4°C. We washed the slices three times before incubating them in secondary antibodies (goat anti-chicken, labeled with Alexa Fluor 488, 1:500, A11039, Invitrogen). We eventually mounted slices with an antifade mounting medium containing DAPI (H-1200, Vectashield). The slices were imaged with a fluorescence microscope (BZ-X710, Keyence; 10 × Plan APO lens) and a confocal microscope (LSM 800, Carl Zeiss; 10 × Plan APO lens). Images were then processed with Fiji software. Fluorescence intensities were modified for visualization purposes.

Behavioral analysis—To determine whether mice learned the task after training, we calculated the sensitivity index, d', that measures discriminability,[46] for each mouse on each training day by: $d' = z(\mathrm{LCR}) - z(\mathrm{HWR})$. Here, z refers to the normal inverse cumulative distribution function while LCR and HWR stand for low-frequency-trial correct-choice rate and high-frequency-trial wrong-choice rate, respectively. LCR was calculated by dividing the number of all LC trials by the sum of all LC and LW trials. Similarly, HWR was calculated

by dividing the number of all HW trials by the sum of all HW and HC trials. The sensitivity index will be the same if it is alternatively calculated by: $d' = z(HCR) - z(LWR)$, where HCR and LWR stand for high-frequency-trial correct-choice rate and low-frequency-trial wrong-choice rate, respectively. In most of our data analysis, we divided the whole training period of 25 days into five stages, with 5 days in each to reduce potential variabilities across training days and to ensure a sufficient number of trials (~500) in each stage for model training. We did not unevenly divide the training stages based on the performance of each animal as it is more subjective and could introduce bias to our results. In addition, the results would also not be comparable with uneven time bins. Moreover, from the performance plot (Figure 1C), we can see that the differences among mice in terms of d' are just within days. Such small variations will not affect our interpretation of the results.

Pre-processing of calcium fluorescence data—Widefield imaging data $F(x, y, t)$ were processed to acquire ΔF/F data by choosing the average intensity of 1 s (30 frames) before the onset of each tone (stimulus) i (at time t_i) as baseline intensity: $F_{i,0}(x, y) = \overline{F}(x, y, [t_i - 1, t_i))$ and computed ΔF/F of each frame of interest j (at time t_j) of stimulus i: $F'(x, y, t_j) = (F(x, y, t_j) - F_{i,0}(x, y))/F_{i,0}(x, y)$. To reduce computational burden and eliminate potential vascular artifacts, imaging data were then spatially down-sampled to have a pixel distance of 75 μm, and a 3×3 median filter was applied. Displayed intensity of average ΔF/F map, $I(x, y)$, was then obtained by normalizing average ΔF/F data with: $I(x, y) = \max(\min F'(x, y, \overline{t})/I_{Max}, 1), 0)$, where I_{Max} was subject to change in different occasions.

Dividing of subregions in the auditory cortex—Subregions of ACtx were determined based on tone-related cortical responses as revealed by ΔF/F changes. Miss trials from the first training day were averaged from 0 to 1 s after the tone onset to generate average ΔF/F maps for low-frequency tone response (Figure S4A, left) and high-frequency tone response (Figure S4A, right) accordingly. A 3 × 3 mean filter was applied on each of the average ΔF/F maps, and local maximum points were searched on each average ΔF/F map. Only local maxima points with ΔF/F value higher than a specific percentile (75% in most cases) of all ΔF/F values within each average ΔF/F map were considered as tone-related response centers for later analysis. Local maxima points in both the average low-frequency response map (Figure S4B, blue points with intensity corresponding to ΔF/F value at each point) and the average high-frequency response map (Figure S4B, red points with intensity corresponding to ΔF/F value at each point) were then combined as expanding centers for each group of tone response region in the following analysis to include all tone-related response areas in each frequency. During each expansion step, vertically and horizontally adjacent pixels for each group were considered candidates to be included in that group. For each pixel, an index was calculated by dividing the ΔF/F value of that pixel by the ΔF/F value of the center point of the group, and a threshold was then chosen such that any pixel with an index value higher than the threshold was considered tone response pixel of that group, whereas those failed to meet the criteria were dropped out. The threshold was determined by $50\% \times \frac{c_i}{c_{max}}$, where c_i is the ΔF/F value of the center point of the group, and c_{max} is the maximum ΔF/F value among all group centers of each corresponding tone frequency. Pixels that met inclusion criteria for more than one group were included in the group with the

highest index value. Expansion steps were repeated five times or until none of the expanding pixels met inclusion criteria during a step, whichever came earlier. Subregions of the ACtx were then manually determined from expanded groups (Figure S4C, blue, red and green groups) based on the frequency response areas map[48,76] and anatomical positions, including low-frequency response areas in the primary auditory cortex ($A1_L$), high-frequency response areas in the primary auditory cortex ($A1_H$), low-frequency response areas in the secondary auditory cortex ($A2_L$), high-frequency response areas in the secondary auditory cortex ($A2_H$), low-frequency response areas in the anterior auditory field (AAF_L), high-frequency response areas in the anterior auditory field (AAF_H), the dorsal medial auditory field (DM), the dorsal anterior auditory field (DA), and the dorsal posterior auditory field (DP). Regions not responding to tone stimuli but were within the minimum convex enclosing polygon of all tone response regions were then processed by convolving the subtraction of all tone response regions from the area of the polygon with a 3 × 3 all-ones kernel and by taking the inverse of convolved image. Non-response dorsal anterior auditory field (DA_n), and non-response dorsal posterior auditory field (DP_n) were then separated manually (Figure S4C, gray groups).

Qualification of response asymmetricity—To quantify the response discrepancy for two stimuli, we calculated the average absolute difference between the mean ΔF/F traces of LC and HC trials during each time interval (the hold or the move phase), for each animal, in each ACtx subregion, and on each training day. We smoothed the results by a median filter of size 5 and calculated the z-scores to normalize the responses relative to the baseline (1 s before the tone onset).

Calculation of inter-regional correlation—We calculated the Pearson correlation coefficient of determination for the average ΔF/F traces between any two subregions in ACtx for each trial at each training stage and during each phase.

Models for decoding cortical responses—Models for decoding presented tones based on responses from the auditory cortex were constructed based on Long Short-Term Memory (LSTM) networks.[39] For the training of each model, imaging sequence data were flattened for each trial as a two-dimensional matrix with columns representing ΔF/F values at each time point, and rows standing for each pixel within the region of interest in the auditory cortex. Long Short-Term Memory (LSTM) network was used as a bi-direction LSTM layer in the network with 250 cells to train data from the input layer and was connected to a dropout layer with a dropout probability of 50%, a fully connected layer, a softmax layer and eventually the output layer. The model was trained using an adaptive moment estimation (Adam) optimizer with a gradient threshold of 2, maximum epochs of 50, minimum batch size of 32, and initial learning rate of 10^{-4} and with data shuffled every epoch. The above hyperparameters were tuned to minimize the difference between model performance results on the training dataset and validation dataset while maintaining high model performance. All training parameters were then kept the same for all models. The training dataset, validation dataset, and testing dataset were split by 70%, 20%, and 10% ratios. Leave-one-out cross-validation was performed during the training of each model to evaluate its performance, such that all data would have been used as a testing dataset

after 10-fold cross-validation. The area under the receiver operating characteristic (ROC) curve[50,51] (AUC) was calculated as an evaluation metric after each validation. AUCs of 10 validations were reported for each model for performance comparison. In our analysis of the representation differences in different behavioral outcomes, we first trained the model with matched sample sizes between correct and wrong trials. Because we only randomly down-sampled the larger dataset if the sample sizes were not matched, we did not use all the data in this analysis. We used 5-fold leave-one-out cross-validation here because the number of wrong trials was becoming increasingly low as the training progressed, and we wanted to ensure that during each validation process, datapoints from all classes were presented in the testing dataset. We then trained the model with data only from the last training stage correct trials and validated the model with data from all training stages and behavioral outcomes separately to interpret representation differences. We did not train the model with both correct and wrong trials because the kept-dropping ratio of the number of wrong trials to correct trials as the training progress might reduce the reliability of the models and affect our interpretation of the results. We only used a model trained with data from the last training stages because the representation of training-related stimulus in correct trials was enhanced to the utmost in the last training stage in all ACtx subregions and the results were more reliable when the performance of the model being used was high. Here, we used longer time interval (1/3 s) than the one we used before (1/6 s), because with more data frames being included in the training, the models can be more robust and yield more stable results. We did not use even longer time-intervals because we still wanted to keep the data from the hold and the move phases separated. Similarly, to decode responses from the auditory cortex for predicting behavioral outcomes, the same network as constructed above was adopted for training. Correct trials and wrong trials were being predicted from trials of all frequencies. Since the number of correct trials and the number of wrong trials were becoming increasingly imbalanced as the training stage advanced, we also used 5-fold leave-one-out cross-validation to ensure that during each validation process, datapoints from all classes were presented in the testing dataset and that AUCs could be calculated as evaluation metrics for the performances of models. No tunings of hyperparameters were involved in the model training process, so the training and testing datasets were split by 80% and 20% ratio.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data are presented as mean ± 95% confidence intervals for calcium traces, and mean ± SEM for all other data, if applicable. All statistical tests were performed in MATLAB R2021b. The Kolmogorov-Smirnov test was first used on each data group to see if the data followed a normal distribution. For groups of data that all followed a normal distribution, a two-tailed Student’s t-test was used to test whether the means of the two groups were equal. Should data from any group fail to follow a normal distribution, a two-sided Wilcoxon rank-sum test was performed. The Friedman test was adopted if multiple groups were involved during comparison, and the Tukey’s test was used as a post hoc test for multiple comparisons. For comparisons involving two categorical variables, two-way ANOVA was used. Fisher’s exact test was performed on behavioral results to determine the discriminability of each mouse in distinguishing different tone frequencies. Pearson correlation coefficient was calculated for average ΔF/F traces between any two subregions in the auditory cortex at each training

stage for low-frequency trials or high-frequency trials. All statistical data can be found in the supplementary tables (Tables S1–S6). All *N* values represent the number of mice in each group. Annotations of *p*-values were made by: ***$p < 0.001$, **0.001 % $p < 0.01$, and *0.01 % $p < 0.05$.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Dr. Ji Liu, Dr. Hijee Kang, Chih-ting Chen, Yunru Chen, and Maggie Lowman for technical help, Katherine Maximov for guidance in behavioral training, and Dr. Minzi Chang for help with immunohistochemistry. We also thank Jonah Mittelstadt, Katherine Maximov, Marina Cardoso de Oliveira, Christopher Workman, Chih-ting Chen, Liang Xiang, and Jade Daher for comments on the paper. This work was supported by NIH R01DC017785 (P.O.K.) and R21MH116450 (P.O.K.).

REFERENCES

1. Irvine DRF (2018). Plasticity in the auditory system. Hear. Res 362, 61–73. 10.1016/j.heares.2017.10.011. [PubMed: 29126650]
2. Irvine DRF (2018). Auditory perceptual learning and changes in the conceptualization of auditory cortex. Hear. Res 366, 3–16. 10.1016/j.heares.2018.03.011. [PubMed: 29551308]
3. Yin P, Fritz JB, and Shamma SA (2014). Rapid spectrotemporal plasticity in primary auditory cortex during behavior. J. Neurosci 34, 4396–4408. 10.1523/JNEUROSCI.2799-13.2014. [PubMed: 24647959]
4. Schoups A, Vogels R, Qian N, and Orban G (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature 412, 549–553. 10.1038/35087601. [PubMed: 11484056]
5. Rutkowski RG, and Weinberger NM (2005). Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. Proc. Natl. Acad. Sci. USA 102, 13664–13669. 10.1073/pnas.0506838102. [PubMed: 16174754]
6. Gdalyahu A, Tring E, Polack P-O, Gruver R, Golshani P, Fanselow MS, Silva AJ, and Trachtenberg JT (2012). Associative Fear Learning Enhances Sparse Network Coding in Primary Sensory Cortex. Neuron 75, 121–132. 10.1016/j.neuron.2012.04.035. [PubMed: 22794266]
7. Poort J, Wilmes KA, Blot A, Chadwick A, Sahani M, Clopath C, Mrsic-Flogel TD, Hofer SB, and Khan AG (2022). Learning and attention increase visual response selectivity through distinct mechanisms. Neuron 110, 686–697.e6. 10.1016/j.neuron.2021.11.016. [PubMed: 34906356]
8. Jurjut O, Georgieva P, Busse L, and Katzner S (2017). Learning Enhances Sensory Processing in Mouse V1 before Improving Behavior. J. Neurosci 37, 6460–6474. 10.1523/JNEUROSCI.3485-16.2017. [PubMed: 28559381]
9. Han X, Xu J, Chang S, Keniston L, and Yu L (2022). Multisensory-Guided Associative Learning Enhances Multisensory Representation in Primary Auditory Cortex. Cerebr. Cortex 32, 1040–1054. 10.1093/cercor/bhab264.
10. Pardi MB, Vogenstahl J, Dalmay T, Spanò T, Pu D-L, Naumann LB, Kretschmer F, Sprekeler H, and Letzkus JJ (2020). A thalamocortical top-down circuit for associative memory. Science 370, 844–848. 10.1126/science.abc2399. [PubMed: 33184213]
11. Letzkus JJ, Wolff SBE, Meyer EMM, Tovote P, Courtin J, Herry C, and Lüthi A (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. Nature 480, 331–335. 10.1038/nature10674. [PubMed: 22158104]
12. Polley DB, Heiser MA, Blake DT, Schreiner CE, and Merzenich MM (2004). Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. Proc. Natl. Acad. Sci. USA 101, 16351–16356. 10.1073/pnas.0407586101. [PubMed: 15534214]

13. Gaffan D (1996). Associative and perceptual learning and the concept of memory systems. Brain Res. Cogn. Brain Res 5, 69–80. 10.1016/s0926-6410(96)00042-0. [PubMed: 9049072]

14. Francis NA, Mukherjee S, Koçillari L, Panzeri S, Babadi B, and Kanold PO (2022). Sequential transmission of task-relevant information in cortical neuronal networks. Cell Rep 39, 110878. 10.1016/j.celrep.2022.110878. [PubMed: 35649366]

15. Francis NA, Elgueda D, Englitz B, Fritz JB, and Shamma SA (2018). Laminar profile of task-related plasticity in ferret primary auditory cortex. Sci. Rep 8, 16375. 10.1038/s41598-018-34739-3. [PubMed: 30401927]

16. David SV, and Shamma SA (2013). Integration over multiple timescales in primary auditory cortex. J. Neurosci 33, 19154–19166. 10.1523/JNEUROSCI.2270-13.2013. [PubMed: 24305812]

17. David SV, Fritz JB, and Shamma SA (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. Proc. Natl. Acad. Sci. USA 109, 2144–2149. 10.1073/pnas.1117717109. [PubMed: 22308415]

18. Atiani S, Elhilali M, David SV, Fritz JB, and Shamma SA (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron 61, 467–480. 10.1016/j.neuron.2008.12.027. [PubMed: 19217382]

19. Mesgarani N, David SV, Fritz JB, and Shamma SA (2009). Influence of context and behavior on stimulus reconstruction from neural activity in primary auditory cortex. J. Neurophysiol 102, 3329–3339. 10.1152/jn.91128.2008. [PubMed: 19759321]

20. Fritz J, Shamma S, Elhilali M, and Klein D (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat. Neurosci 6, 1216–1223. 10.1038/nn1141. [PubMed: 14583754]

21. Cook JR, Li H, Nguyen B, Huang H-H, Mahdavian P, Kirchgessner MA, Strassmann P, Engelhardt M, Callaway EM, and Jin X (2022). Secondary auditory cortex mediates a sensorimotor mechanism for action timing. Nat. Neurosci 25, 330–344. 10.1038/s41593-022-01025-5. [PubMed: 35260862]

22. Wang Z-Q, Wen H-Z, Luo T-T, Chen P-H, Zhao Y-D, Wu G-Y, and Xiong Y (2023). Corticostriatal Neurons in the Anterior Auditory Field Regulate Frequency Discrimination Behavior. Neurosci. Bull 39, 962–972. 10.1007/s12264-022-01015-4. [PubMed: 36629979]

23. Yin P, Strait DL, Radtke-Schuller S, Fritz JB, and Shamma SA (2020). Dynamics and Hierarchical Encoding of Non-compact Acoustic Categories in Auditory and Frontal Cortex. Curr. Biol 30, 1649–1663.e5. 10.1016/j.cub.2020.02.047. [PubMed: 32220317]

24. Elgueda D, Duque D, Radtke-Schuller S, Yin P, David SV, Shamma SA, and Fritz JB (2019). State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex. Nat. Neurosci 22, 447–459. 10.1038/s41593-018-0317-8. [PubMed: 30692690]

25. Atiani S, David SV, Elgueda D, Locastro M, Radtke-Schuller S, Shamma SA, and Fritz JB (2014). Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. Neuron 82, 486–499. 10.1016/j.neuron.2014.02.029. [PubMed: 24742467]

26. Dong C, Qin L, Zhao Z, Zhong R, and Sato Y (2013). Behavioral modulation of neural encoding of click-trains in the primary and nonprimary auditory cortex of cats. J. Neurosci 33, 13126–13137. 10.1523/JNEUROSCI.1724-13.2013. [PubMed: 23926266]

27. KASTNER S, and PINSK MA (2004). Visual attention as a multilevel selection process. Cognit. Affect Behav. Neurosci 4, 483–500. 10.3758/CABN.4.4.483. [PubMed: 15849892]

28. Maunsell JHR, and Cook EP (2002). The role of attention in visual processing. Philos. Trans. R. Soc. Lond. B Biol. Sci 357, 1063–1072. 10.1098/rstb.2002.1107. [PubMed: 12217174]

29. Stoilova VV, Knauer B, Berg S, Rieber E, Jäkel F, and Stüttgen MC (2020). Auditory cortex reflects goal-directed movement but is not necessary for behavioral adaptation in sound-cued reward tracking. J. Neurophysiol 124, 1056–1071. 10.1152/jn.00736.2019. [PubMed: 32845769]

30. Liu J, and Kanold PO (2022). Interactive auditory task reveals complex sensory-action integration in mouse primary auditory cortex. Preprint at bioRxiv. 10.1101/2022.12.12.520155.

31. Vivaldo CA, Lee J, Shorkey M, Keerthy A, and Rothschild G (2023). Auditory cortex ensembles jointly encode sound and locomotion speed to support sound perception during movement. PLoS Biol 21, e3002277. 10.1371/journal.pbio.3002277. [PubMed: 37651461]

32. Morandell K, Yin A, Rio RTD, and Schneider DM (2024). Movement-related modulation in mouse auditory cortex is widespread yet locally diverse. J. Neurosci 10.1523/JNEUROSCI.1227-23.2024.
33. Nelson A, Schneider DM, Takatoh J, Sakurai K, Wang F, and Mooney R (2013). A circuit for motor cortical modulation of auditory cortical activity. J. Neurosci 33, 14342–14353. 10.1523/JNEUROSCI.2275-13.2013. [PubMed: 24005287]
34. Yavorska I, and Wehr M (2021). Effects of Locomotion in Auditory Cortex Are Not Mediated by the VIP Network. Front. Neural Circ 15, 618881. 10.3389/fncir.2021.618881.
35. Pienkowski M, and Eggermont JJ (2011). Cortical tonotopic map plasticity and behavior. Neurosci. Biobehav. Rev 35, 2117–2128. 10.1016/j.neubiorev.2011.02.002. [PubMed: 21315757]
36. Takamiya S, Shiotani K, Ohnuki T, Osako Y, Tanisumi Y, Yuki S, Manabe H, Hirokawa J, and Sakurai Y (2022). Auditory Cortex Neurons Show Task-Related and Learning-Dependent Selectivity toward Sensory Input and Reward during the Learning Process of an Associative Memory Task. eNeuro 9. 10.1523/ENEURO.0046-22.2022.
37. Weinberger NM (2004). Specific long-term memory traces in primary auditory cortex. Nat. Rev. Neurosci 5, 279–290. 10.1038/nrn1366. [PubMed: 15034553]
38. Liu J, Maximov K, and Kanold PO (2022). Automated head-fixation training system with high levels of animal participation in psychoacoustic tasks. Preprint at bioRxiv. 10.1101/2022.11.29.518279.
39. Tseng P-H, Urpi NA, Lebedev M, and Nicolelis M (2019). Decoding Movements from Cortical Ensemble Activity Using a Long Short-Term Memory Recurrent Network. Neural Comput 31, 1085–1113. 10.1162/neco_a_01189. [PubMed: 30979355]
40. Georgiou P, Zanos P, Mou T-CM, An X, Gerhard DM, Dryanovski DI, Potter LE, Highland JN, Jenne CE, Stewart BW, et al. (2022). Experimenters' sex modulates mouse behaviors and neural responses to ketamine via corticotropin releasing factor. Nat. Neurosci 25, 1191–1200. 10.1038/s41593-022-01146-x. [PubMed: 36042309]
41. Sorge RE, Martin LJ, Isbester KA, Sotocinal SG, Rosen S, Tuttle AH, Wieskopf JS, Acland EL, Dokova A, Kadoura B, et al. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat. Methods 11, 629–632. 10.1038/nmeth.2935. [PubMed: 24776635]
42. Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300. 10.1038/nature12354. [PubMed: 23868258]
43. Bowen Z, Winkowski DE, and Kanold PO (2020). Functional organization of mouse primary auditory cortex in adult C57BL/6 and F1 (CBAxC57) mice. Sci. Rep 10, 10905. 10.1038/s41598-020-67819-4. [PubMed: 32616766]
44. Francis NA, Winkowski DE, Sheikhattar A, Armengol K, Babadi B, and Kanold PO (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. Neuron 97, 885–897.e6. 10.1016/j.neuron.2018.01.019. [PubMed: 29398362]
45. Bimbard C, Sit TPH, Lebedeva A, Reddy CB, Harris KD, and Carandini M (2023). Behavioral origin of sound-evoked activity in mouse visual cortex. Nat. Neurosci 26, 251–258. 10.1038/s41593-022-01227-x. [PubMed: 36624279]
46. Nevin JA (1969). SIGNAL DETECTION THEORY AND OPERANT BEHAVIOR: A Review of David M. Green and John A. Swets' Signal Detection Theory and Psychophysics. [1]. J. Exp. Anal. Behav 12, 475–480. 10.1901/jeab.1969.12-475.
47. Liu J, Whiteway MR, Sheikhattar A, Butts DA, Babadi B, and Kanold PO (2019). Parallel Processing of Sound Dynamics across Mouse Auditory Cortex via Spatially Patterned Thalamic Inputs and Distinct Areal Intracortical Circuits. Cell Rep 27, 872–885.e7. 10.1016/j.celrep.2019.03.069. [PubMed: 30995483]
48. Tsukano H, Horie M, Ohga S, Takahashi K, Kubota Y, Hishida R, Takebayashi H, and Shibuki K (2017). Reconsidering Tonotopic Maps in the Auditory Cortex and Lemniscal Auditory Thalamus in Mice. Front. Neural Circ 11, 14. 10.3389/fncir.2017.00014.
49. Graves A, and Schmidhuber J (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Network. 18, 602–610. 10.1016/j.neunet.2005.06.042.

50. Fawcett T (2006). An introduction to ROC analysis. Pattern Recogn. Lett 27, 861–874. 10.1016/j.patrec.2005.10.010.
51. Huang J, and Ling CX (2005). Using AUC and accuracy in evaluating learning algorithms. IEEE Trans. Knowl. Data Eng 17, 299–310. 10.1109/TKDE.2005.50.
52. Manns M, Basbasse YE, Freund N, and Ocklenburg S (2021). Paw preferences in mice and rats: Meta-analysis. Neurosci. Biobehav. Rev 127, 593–606. 10.1016/j.neubiorev.2021.05.011. [PubMed: 34004244]
53. Murphy TH, Michelson NJ, Boyd JD, Fong T, Bolanos LA, Bierbrauer D, Siu T, Balbi M, Bolanos F, Vanni M, and LeDue JM (2020). Automated task training and longitudinal monitoring of mouse mesoscale cortical circuits using home cages. Elife 9, e55964. 10.7554/eLife.55964. [PubMed: 32412409]
54. Murphy TH, Boyd JD, Bolaños F, Vanni MP, Silasi G, Haupt D, and LeDue JM (2016). High-throughput automated home-cage mesoscopic functional imaging of mouse cortex. Nat. Commun 7, 11611. 10.1038/ncomms11611. [PubMed: 27291514]
55. Niekisch H, Steinhardt J, Berghäuser J, Bertazzoni S, Kaschinski E, Kasper J, Kisse M, Mitlöhner J, Singh JB, Weber J, et al. (2019). Learning Induces Transient Upregulation of Brevican in the Auditory Cortex during Consolidation of Long-Term Memories. J. Neurosci 39, 7049–7060. 10.1523/JNEUROSCI.2499-18.2019. [PubMed: 31217331]
56. Schicknick H, Reichenbach N, Smalla K-H, Scheich H, Gundelfinger ED, and Tischmeyer W (2012). Dopamine modulates memory consolidation of discrimination learning in the auditory cortex. Eur. J. Neurosci 35, 763–774. 10.1111/j.1460-9568.2012.07994.x. [PubMed: 22339853]
57. Adibi M, McDonald JS, Clifford CWG, and Arabzadeh E (2013). Adaptation improves neural coding efficiency despite increasing correlations in variability. J. Neurosci 33, 2108–2120. 10.1523/JNEUROSCI.3449-12.2013. [PubMed: 23365247]
58. von der Behrens W, Bäuerle P, Kössl M, and Gaese BH (2009). Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. J. Neurosci 29, 13837–13849. 10.1523/JNEUROSCI.3475-09.2009. [PubMed: 19889995]
59. Nelken I, and Ulanovsky N (2007). Mismatch Negativity and Stimulus-Specific Adaptation in Animal Models. J. Psychophysiol 21, 214–223. 10.1027/0269-8803.21.34.214.
60. Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, and Kilgard MP (2011). Cortical map plasticity improves learning but is not necessary for improved performance. Neuron 70, 121–131. 10.1016/j.neuron.2011.02.038. [PubMed: 21482361]
61. Polley DB, Steinberg EE, and Merzenich MM (2006). Perceptual Learning Directs Auditory Cortical Map Reorganization through Top-Down Influences. J. Neurosci 26, 4970–4982. 10.1523/JNEUROSCI.3771-05.2006. [PubMed: 16672673]
62. Caras ML, and Sanes DH (2017). Top-down modulation of sensory cortex gates perceptual learning. Proc. Natl. Acad. Sci. USA 114, 9972–9977. 10.1073/pnas.1712305114. [PubMed: 28847938]
63. Hochstein S, and Ahissar M (2002). View from the top: hierarchies and reverse hierarchies in the visual system. Neuron 36, 791–804. 10.1016/s0896-6273(02)01091-7. [PubMed: 12467584]
64. Hackett TA (2011). Information flow in the auditory cortical network. Hear. Res 271, 133–146. 10.1016/j.heares.2010.01.011. [PubMed: 20116421]
65. Grosso A, Cambiaghi M, Renna A, Milano L, Roberto Merlo G, Sacco T, and Sacchetti B (2015). The higher order auditory cortex is involved in the assignment of affective value to sensory stimuli. Nat. Commun 6, 8886. 10.1038/ncomms9886. [PubMed: 26619940]
66. Huang Y, Matysiak A, Heil P, König R, and Brosch M (2016). Persistent neural activity in auditory cortex is related to auditory working memory in humans and nonhuman primates. Elife 5, e15441. 10.7554/eLife.15441. [PubMed: 27438411]
67. Cambiaghi M, Renna A, Milano L, and Sacchetti B (2017). Reversible Inactivation of the Higher Order Auditory Cortex during Fear Memory Consolidation Prevents Memory-Related Activity in the Basolateral Amygdala during Remote Memory Retrieval. Front. Behav. Neurosci 11, 138. 10.3389/fnbeh.2017.00138. [PubMed: 28790901]

68. Li R, Huang J, Li L, Zhao Z, Liang S, Liang S, Wang M, Liao X, Lyu J, Zhou Z, et al. (2023). Holistic bursting cells store long-term memory in auditory cortex. Nat. Commun 14, 8090. 10.1038/s41467-023-43620-5. [PubMed: 38062015]

69. Mohn JL, Baese-Berk MM, and Jaramillo S (2024). Selectivity to acoustic features of human speech in the auditory cortex of the mouse. Hear. Res 441, 108920. 10.1016/j.heares.2023.108920. [PubMed: 38029503]

70. Zaehle T, Geiser E, Alter K, Jancke L, and Meyer M (2008). Segmental processing in the human auditory dorsal stream. Brain Res 1220, 179–190. 10.1016/j.brainres.2007.11.013. [PubMed: 18096139]

71. Rogalsky C, Basilakos A, Rorden C, Pillay S, LaCroix AN, Keator L, Mickelsen S, Anderson SW, Love T, Fridriksson J, et al. (2022). The Neuroanatomy of Speech Processing: A Large-scale Lesion Study. J. Cognit. Neurosci 34, 1355–1375. 10.1162/jocn_a_01876. [PubMed: 35640102]

72. Calhoun G, Chen C-T, and Kanold PO (2023). Bilateral widefield calcium imaging reveals circuit asymmetries and lateralized functional activation of the mouse auditory cortex. Proc. Natl. Acad. Sci. USA 120, e2219340120. 10.1073/pnas.2219340120. [PubMed: 37459544]

73. Yasuda SP, Seki Y, Suzuki S, Ohshiba Y, Hou X, Matsuoka K, Wada K, Shitara H, Miyasaka Y, and Kikkawa Y (2020). c.753A>G genome editing of a Cdh23ahl allele delays age-related hearing loss and degeneration of cochlear hair cells in C57BL/6J mice. Hear. Res 389, 107926. 10.1016/j.heares.2020.107926. [PubMed: 32101784]

74. Liu J, and Kanold PO (2021). Diversity of Receptive Fields and Sideband Inhibition with Complex Thalamocortical and Intracortical Origin in L2/3 of Mouse Primary Auditory Cortex. J. Neurosci 41, 3142–3162. 10.1523/JNEUROSCI.1732-20.2021. [PubMed: 33593857]

75. Pnevmatikakis EA, and Giovannucci A (2017). NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. J. Neurosci. Methods 291, 83–94. 10.1016/j.jneumeth.2017.07.031. [PubMed: 28782629]

76. Tsukano H, Horie M, Bo T, Uchimura A, Hishida R, Kudoh M, Takahashi K, Takebayashi H, and Shibuki K (2015). Delineation of a frequency-organized region isolated from the mouse primary auditory cortex. J. Neurophysiol 113, 2900–2920. 10.1152/jn.00932.2014. [PubMed: 25695649]

Highlights

- Auditory learning recruits dorsal posterior (DP) and dorsal anterior (DA) auditory fields
- Neural representation of behavior and learned stimuli is enhanced in DP and DA, respectively
- Functional connectivity between the primary and higher-order auditory regions is enhanced
- Induced by learning, response modulation is also affected by task-associated movement

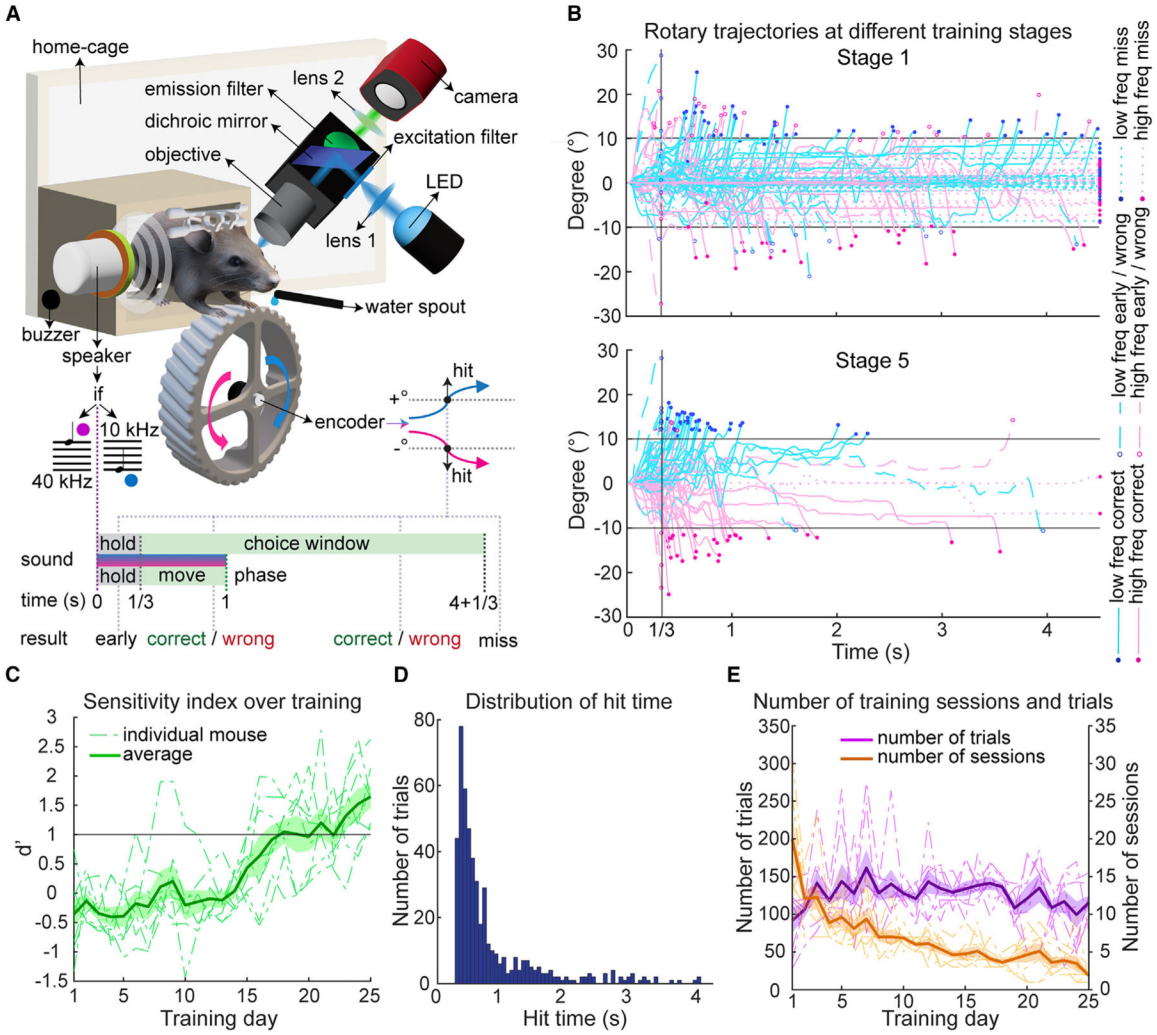


Figure 1. Behavioral performance on an auditory discrimination task with longitudinal widefield imaging

(A) Top, an overview of the behavioral and imaging system. Bottom, the training paradigm of the tone discrimination task. In this task, the animals were required to turn a wheel to the left (positive degrees) when a low-frequency (10 kHz) tone was presented and to the right (negative degrees) when a high-frequency (40 kHz) tone was presented. The choice was considered made once the wheel was turned to pass a threshold of 10° in either direction. Any decision made during the hold period (0–1/3 s from tone onset) was considered an early response trial, where a buzzer was activated as punishment. A decision made during the choice window (4 s after the end of the hold period) was either correct (rewarded with water) or wrong (punished with prolonged timeout duration). A trial with no decision made during the choice window was considered a miss trial. Neural responses from the hold and move phases were used for analysis.

(B) Randomly selected rotary results at the first (top, training days 1–5) and the last (bottom, training days 21–25) training stages of an exemplar mouse. Thresholds are plotted as horizontal lines where a positive degree corresponds to the left turning of the wheel.

(C) Behavioral performances as indicated by the sensitivity index. The shaded area represents the standard error of mean (SEM) of all mice ($N = 8$). Threshold at $d' = 1$ is plotted, which indicates the ability of discrimination. Fisher's exact test results of each mouse on the last training day: ***$p = 1.6e{-}15$, ***$p = 1.7e{-}11$, ***$p = 1.9e{-}8$, ***$p = 3.8e{-}7$, ***$p = 7.9e{-}5$, ***$p = 2.3e{-}4$, **$p = 0.0097$, and *$p = 0.018$.

(D) Histogram showing hit times of all correct trials from all mice ($N = 8$) on the last training day.

(E) Plots showing the number of training sessions and the total trials performed by each mouse on each training day. Dashed lines show results from each mouse, with the shaded area representing the SEM of all mice ($N = 8$).

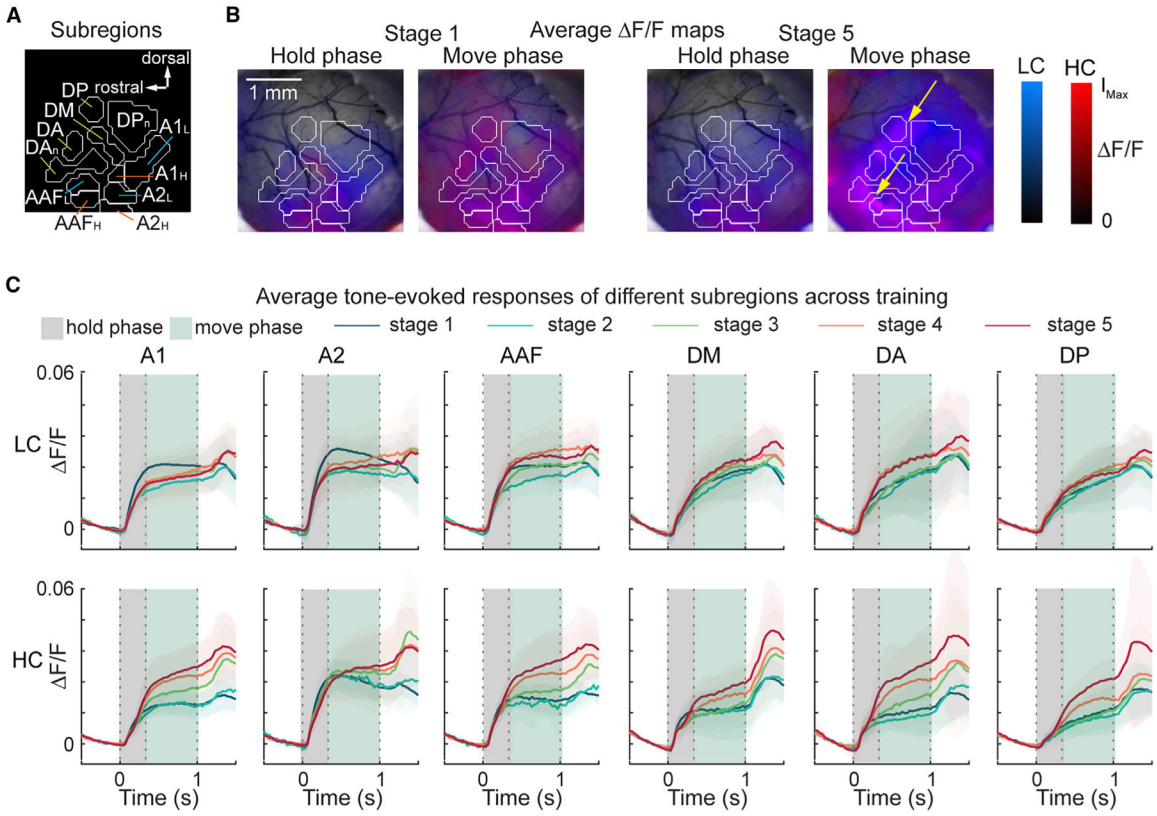


Figure 2. Changes in tone-evoked response magnitudes

(A) Boundaries of the determined subregions in the auditory cortex (ACtx): A1, primary ACtx; $A1_L$, low-frequency response area in A1; $A1_H$, high-frequency response area in ACtx; A2, secondary ACtx; $A2_L$, low-frequency response area in A2; $A2_H$, high-frequency response area in A2; AAF, anterior auditory field; AAF_L, low-frequency response area in AAF; AAF_H, high-frequency response area in AAF; DM, dorsal medial auditory field; DA, dorsal anterior auditory field; DP, dorsal posterior auditory field; DA_n, non-response DA area; DP_n, non-response DP area.

(B) Imaging results of an exemplar mouse at the first and last training stages. Average ΔF/F maps during the hold or move phase for low- or high-frequency correct (LC or HC) trials are shown with intensities normalized by I_{Max} = 2:5% in ΔF/F values for maps during the hold phase and I_{Max} = 4:0% in ΔF/F values for maps during the move phase and are superimposed on the raw widefield images (scale bar, 1 mm).

(C) Average ΔF/F traces of different training stages and ACtx subregions. The shaded area represents the 95% confidence interval of all mice (for each trace in A1: N = 8, A2: N = 6, AAF: N = 7, DM: N = 8, DA: N = 5, DP: N = 7).

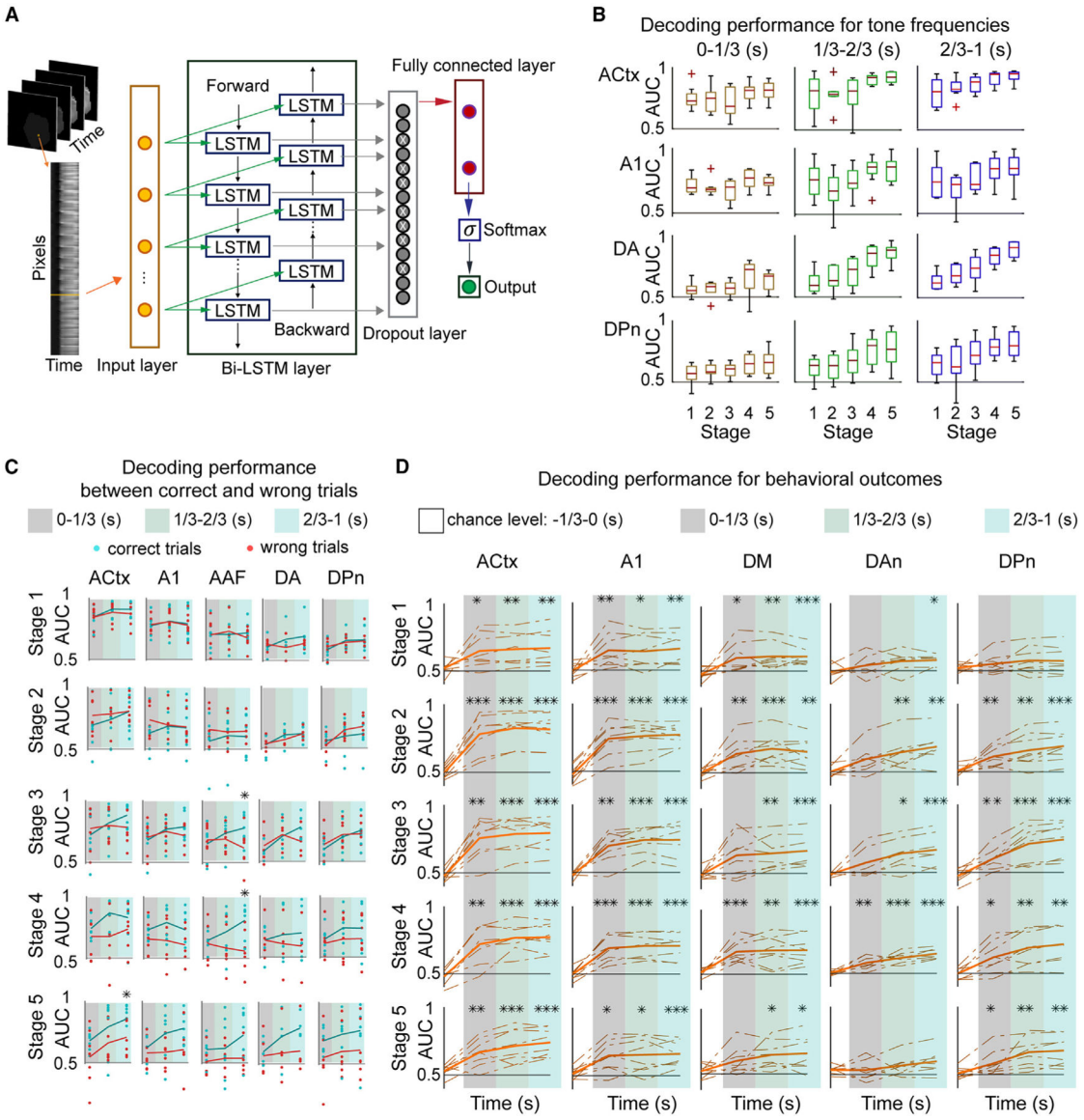


Figure 3. Decoding capabilities for tone frequencies in different ACtx subregions and behavioral contexts and for different behavioral outcomes across training stages

(A) Schema of the long short-term memory network used for the model training. Imaging sequence data were flattened for each trial as a two-dimensional matrix with columns representing ΔF/F values at each time point and rows standing for each pixel in the ACtx or each subregion.

(B) Performance of models trained on various datasets of all correct trials for decoding tone frequencies. Average AUC results from different ACtx subregions are shown for all mice (for each box in ACtx: $N = 8$, A1: $N = 8$, DA: $N = 5$, DPn: $N = 8$) across different training stages for three different time intervals.

(C) Performance of models in predicting presented tone frequencies from correct or wrong trials at different training stages. Solid lines show the average testing results of all mice as a function of different time intervals involved in the model training. Statistical results are shown (Table S2), with $*p < 0.05$.

(D) Decoding performance of behavioral outcomes from different ACtx subregions, training stages, and behavioral phases. For each image, the average AUCs from leave-one-out 5-fold cross-validation for models trained on four data time intervals, namely, from the −1/3 to 0, 0 to 1/3, 1/3 to 2/3, and 2/3 to 1 s periods w.r.t. the tone onset, are presented as model performance results for each mouse as plotted as dashed lines. The average performance results of all mice asa function of different behavioral phases involved in the model training are plotted as solid lines. Statistical results are shown (Table S3), with *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

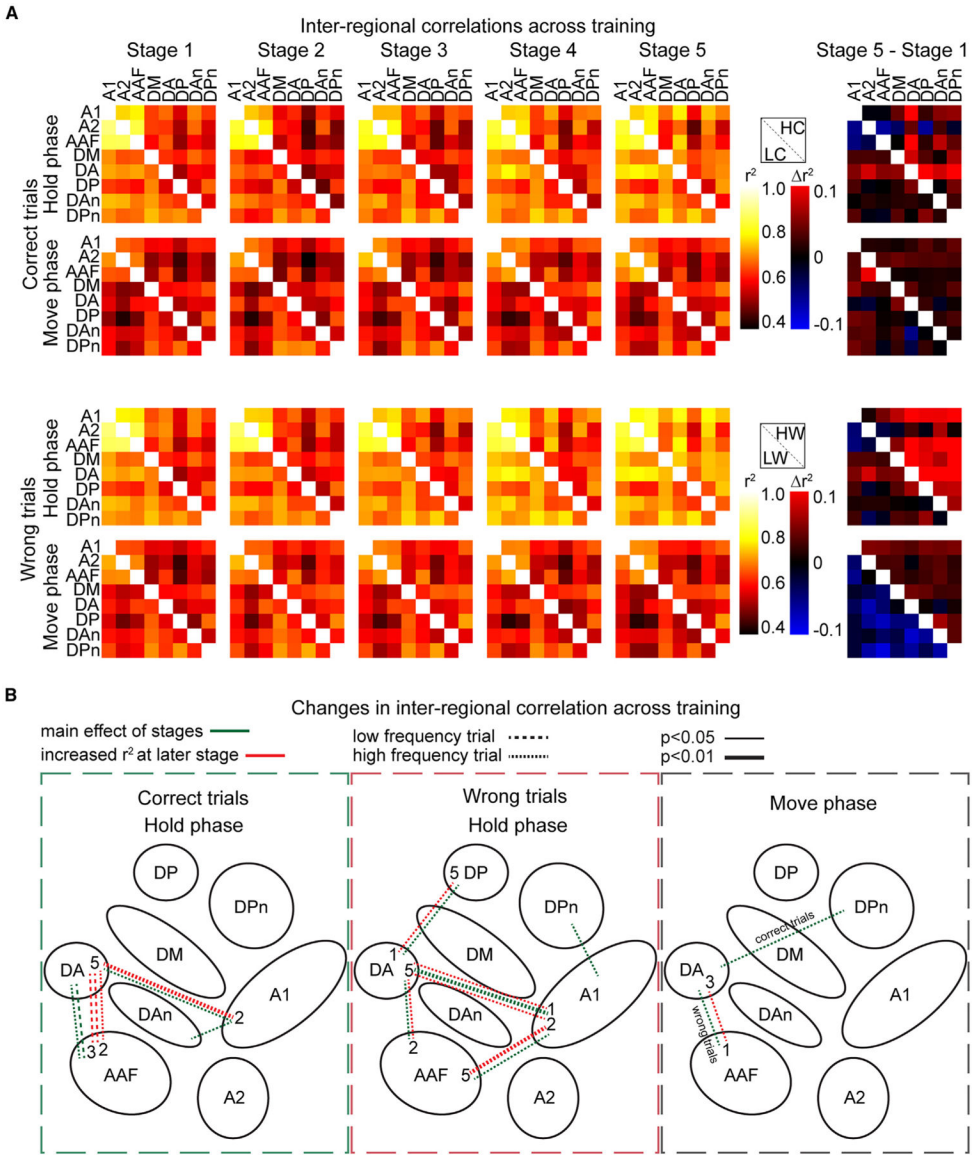


Figure 4. Inter-regional correlations of the tone-related responses from correct or wrong trials over training

(A) Heatmaps showing the average values across mice of the Pearson correlation coefficient of determination for average ΔF/F traces between any two subregions in the ACtx for each trial at each training stage and during each phase. The first two rows are inter-regional correlations from LC or HC trials. The last two rows are inter-regional correlations from low- or high-frequency wrong (LW or HW) trials. Values below 0.4 are set to black in the heatmaps. The differences between correlations from the fifth and the first training stages are plotted to the right.

(B) Diagrams showing the changes of inter-regional correlations among all stages (main effect) or between each two training stages (post hoc multiple comparisons, the number at each end of a line) in correct or wrong trials during each phase (Table S5).

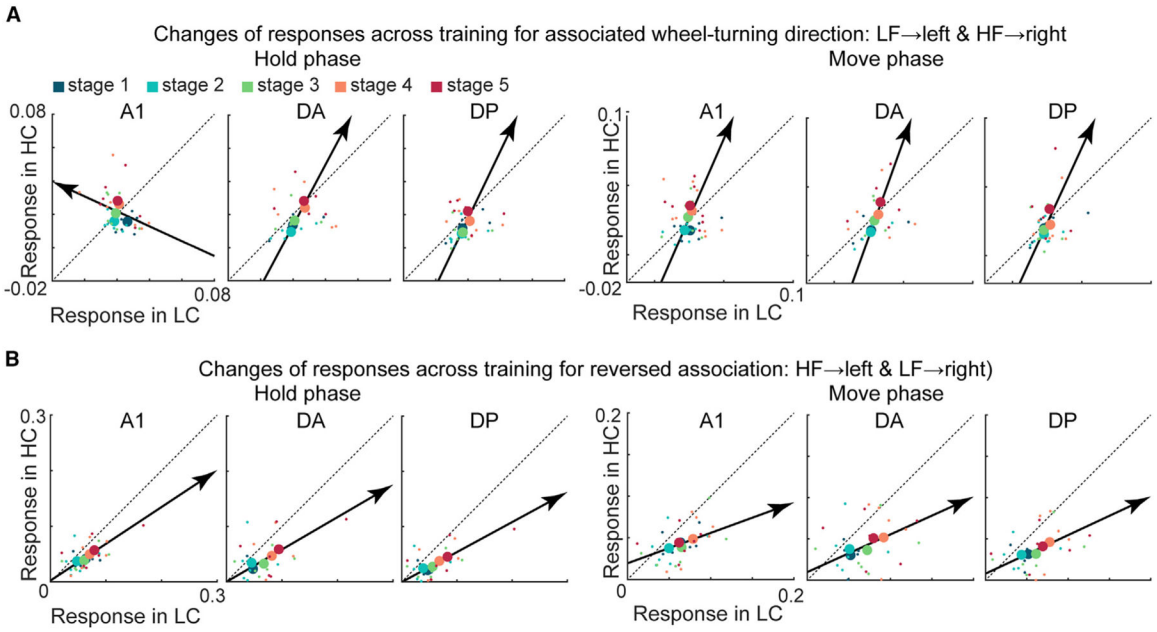


Figure 5. Changes of responses across training for different associated wheel-turning directions

(A and B) To measure the asymmetricity of the response changes, we first calculated the averaged ΔF/F traces from either LC or HC trials and averaged the traces during each time interval (the hold or move phase), for each animal, in each ACtx subregion, and at each training stage. We then obtained different sets of response-character dots for different animals at different training stages for different phases and subregions, where the x attribute represents the average response in LC and the y attribute represents the average response in HC, with small dots corresponding to data from each animal and at each training stage. For each training stage, we then calculated the average response-character dot (larger dots) of all the animals (for A: A1: $N = 8$, DA: $N = 5$, DP: $N = 8$; for B: $N = 6$ in all subregions), and we performed linear regression for the average dots from stages 1–5. In the regression plot, we point in the direction from the dot corresponding to the first training stage toward the dot for the last training stage and regard the vector direction w.r.t. the line y = x (plotted as a dashed line) as the asymmetricity of the response changes.

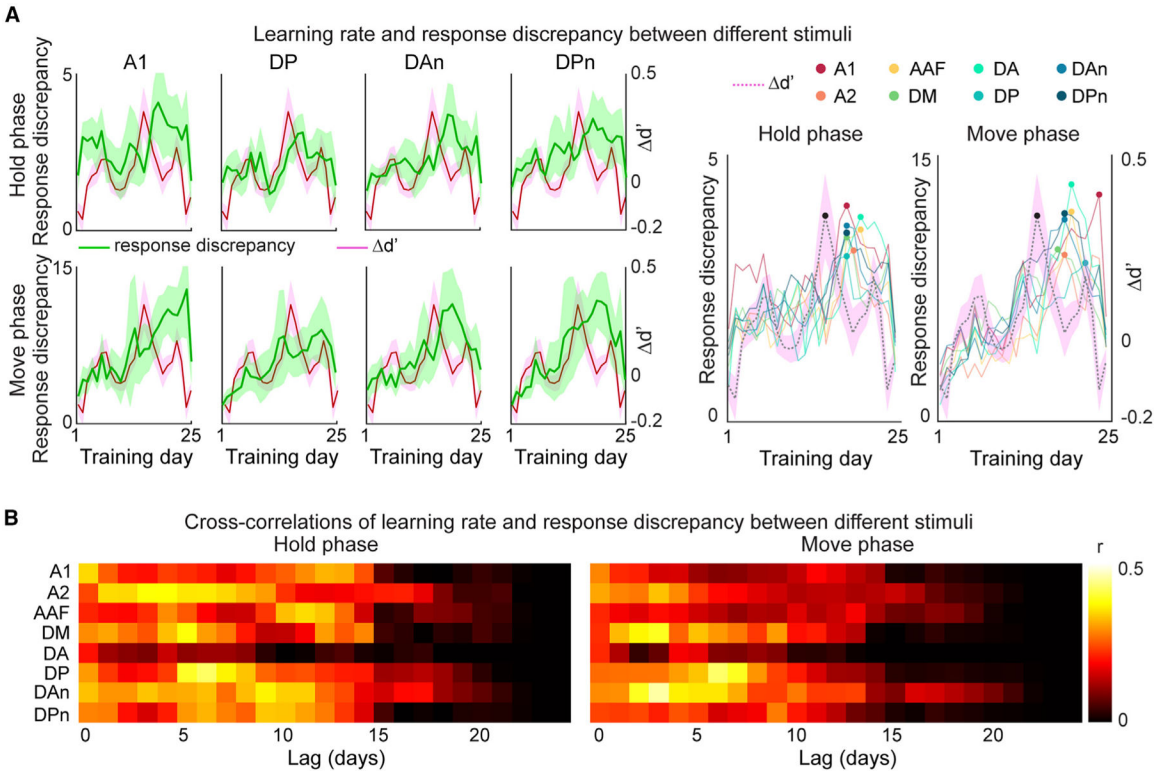


Figure 6. Synchronicity between the learning rate and the response discrepancy

(A) The learning rate and response discrepancy across training. Left, results for selected subregions plotted separately. Right, overlaid results of selected subregions. To quantify the learning rate, we filtered the d′ by a median filter of size 5 and then calculated the derivative of the d′ (Δd′). To quantify the response discrepancy, the average absolute differences between ΔF/F traces from LC trials and HC trials were calculated during each time interval (the hold or move phase), for each animal, in each ACtx subregion, and on each training day. We then filtered the results by a median filter of size 5 and performed the normalization by calculating the *Z* scores using the same values from baseline ΔF/F traces (1 s before the tone onset). Solid lines indicate the average results of all mice, and the shaded areas indicate the SEM of all mice (A1: $N = 8$, DP: $N = 7$, DA_n: $N = 7$, DP_n: $N = 8$, Δ d′: $N = 8$). The peak in each trace in the right images is shown by a solid dot.

(B) Cross-correlation results of the learning rate and the response discrepancy. Cross-correlations were performed for each mouse, and the average results of all mice are shown (A1: $N = 8$, A2: $N = 6$, AAF: $N = 7$, DM: $N = 8$, DA: $N = 5$, DP: $N = 7$, DA_n: $N = 7$, DP_n: $N = 8$).

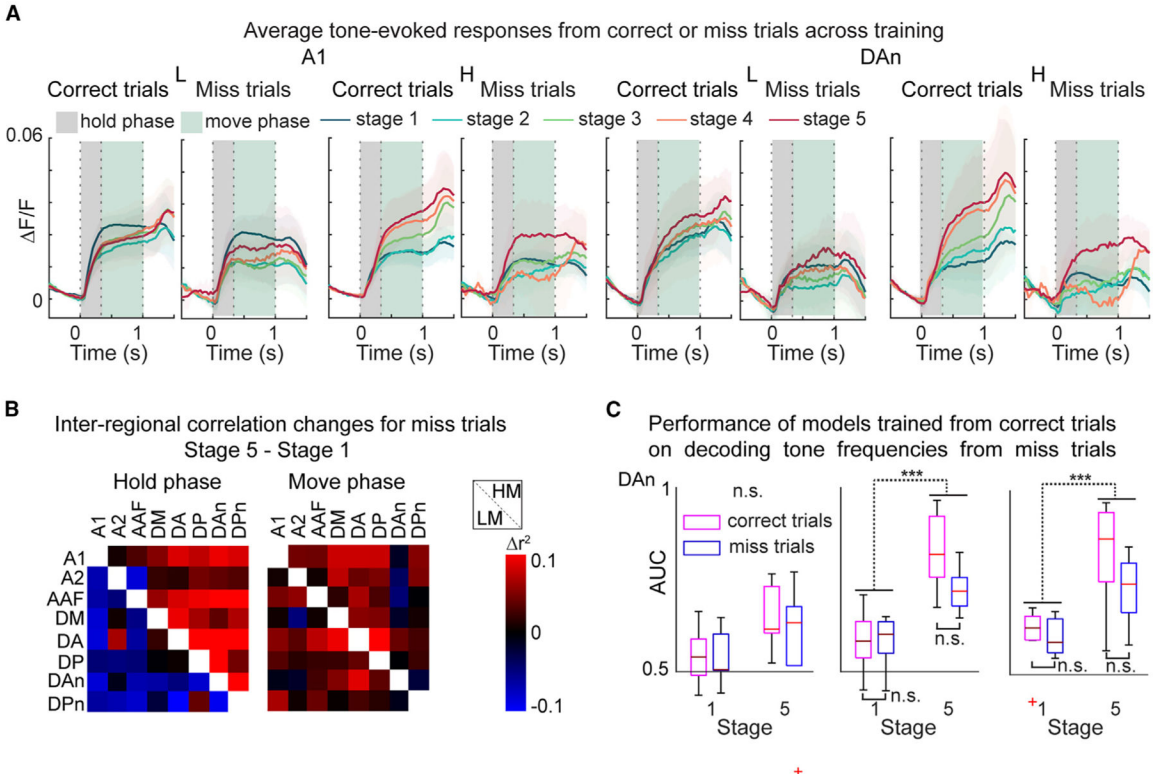


Figure 7. The differences between changes in correct and miss trials

(A) Average ΔF/F traces for different training stages and selected ACtx subregions of correct or miss trials. The shaded area represents the 95% confidence intervals of all mice (for each trace in A1: $N = 8$, DA_n: $N = 7$).

(B) The differences between the Pearson correlation coefficients of determination from the fifth and the first training stages.

(C) Validating models trained from correct trials on miss trials for decoding tone frequencies. Statistical results are shown (Table S6), with $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Calcium imaging data	This paper	https://doi.org/10.7281/T1/J8UWFU
Behavioral data	This paper	https://doi.org/10.7281/T1/J8UWFU
Experimental models: Organisms/strains		
Mouse: C57BL/6J-Tg(Thy1-GCaMP6s)GP4.3Dkim/J	The Jackson Laboratory	JAX# 024275
Mouse: B6.CAST-Cdh23Ahl+/Kjn	The Jackson Laboratory	JAX# 002756
Software and algorithms		
MATLAB	MathWorks	https://www.mathworks.com
Automated training	Liu et al.[38]	https://doi.org/10.1101/2022.11.29.518279.
Automated region of interest matching	This paper	https://doi.org/10.7281/T1/J8UWFU
Motion correction	https://github.com/flatironinstitute/NoRMCorre	https://doi.org/10.1016/j.jneumeth.2017.07.031
Data analysis code	This paper	https://doi.org/10.7281/T1/J8UWFU
Other		
CMOS camera	Thorlabs	CS505MU
blue LED	Thorlabs	M470L5
T-Cube LED Driver	Thorlabs	LEDD1A
Plano-Convex lens	Thorlabs	LA1951-A
excitation filter	Chroma	ET470/40
dichroic mirror	Thorlabs	MD499
Epi-Illuminator module	Thorlabs	WFA2002
objective	Olympus	UPlanFl 4×/0.13
emission filter	Chroma	AT535/40m
camera tube	Thorlabs	WFA4102–0.5×
T-Cube DC Servo Motor Controller	Thorlabs	TDC001
Actuator	Thorlabs	Z812B