

Age and Learning Shapes Sound Representations in Auditory Cortex During Adolescence

Reviewed Preprint

v3 • August 21, 2025

Revised by authors

Reviewed Preprint

v2 • July 15, 2025

Reviewed Preprint

v1 • March 31, 2025

Benedikt Praegel, Feng Chen, Adria Dym, Amichai Lavi-Rudel, Shaul Druckmann, Adi Mizrahi 

The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel • Department of Neurobiology, the Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel • Department of Neurobiology, Stanford University, Stanford, United States • Department of Applied Physics, Stanford University, Stanford, United States • Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, United States

 https://en.wikipedia.org/wiki/Open_access
 Copyright information

eLife Assessment

This **important** study suggests that adolescent mice exhibit less accuracy than adult mice in a sound discrimination task when the sound frequencies are very similar. The evidence supporting this observation is **solid** and suggests that it arises from cognitive control differences between adolescent and adult mice. The adolescent period is largely understudied, despite its contribution to shaping the adult brain, which makes this study interesting for a broad range of neuroscientists.

<https://doi.org/10.7554/eLife.106387.3.sa3>

Abstract

Adolescence is a developmental period characterized by heightened plasticity. Yet, how ongoing development affects sensory processing and cognitive function is unclear. We investigated how adolescent (postnatal day 20-42) and adult (postnatal day 60-82) mice differ in performance on a pure tone Go/No-Go auditory discrimination task of varying difficulty. Using dense electrophysiological recordings, we measured spiking activity at single neuron resolution in the auditory cortex while mice were engaged in the task. As compared to adults, adolescent mice showed lower auditory discrimination performance in a difficult task. This difference in performance was due to higher response variability and weaker cognitive control expressed as higher lick bias. Adolescent and adult neuronal responses differed only slightly in representations of pure tones when measured outside the context of learning and the task. However, cortical representations after learning within the context of the task were markedly different. We found differences in stimulus- and choice-related activity at the single neuron level representations, as well as lower population-level decoding of the difficult task in adolescents. Overall, cortical decoding in adolescents was lower and slower, especially for difficult sound discrimination, reflecting immature cortical representations of sounds and choices. Notably, we found age-related differences, which were more pronounced after learning, reflecting the combined impact of age and learning. Our findings highlight distinct neurophysiological and behavioral profiles in adolescence, underscoring the ongoing

development of cognitive control mechanisms and cortical plasticity during this sensitive developmental period.

Introduction

Adolescence is a developmental stage characterized by the continued refinement of perception and cognition. During this period, the brain is highly sensitive to environmental influences, making it particularly vulnerable to both negative experiences, such as substance abuse, and positive influences, such as supportive relationships[1, 2]. While the adolescent brain exhibits greater plasticity compared to adults[3], the extent to which specific sensory and cognitive traits are fully developed versus those that remain malleable is still debated. For instance, some studies on sensory perception have demonstrated that adolescent learning is slower and more variable[4, 5], while others have found that adolescents learn faster in novel contexts, such as reversal learning[6]. Certain traits, such as cognitive flexibility, are more pronounced in adolescents compared to adults, whereas other behaviors such as impulsive control, are still immature[7]. Furthermore, results across studies that target sensory or cognitive traits are often inconsistent[8]. Therefore, understanding the degree to which sensory and cognitive functions have matured, and the underlying brain mechanisms that support them, remain an ongoing challenge.

A key concept in brain development is that of critical periods (CPs), which are specific time windows during which the brain's neural circuits are highly plastic and particularly sensitive to external experiences[9, 10]. The early postnatal development of sensory cortices provides classical examples of this phenomenon. In mice, neurons in the visual cortex undergo several CPs such as those for orientation selectivity, direction selectivity and ocular dominance. During the CP, which occurs in different developmental time windows for different functional attributes, visual experience shapes the specific neuronal attribute to a long-term state (e.g., ocular dominance reaches its adult form)[9, 11]. The CPs for simple neuronal features in primary sensory cortices are typically complete by the onset of puberty. In the auditory cortex (ACx) of mice, the CP for pure tone processing begins as early as postnatal day 12 (P12), shortly after the ear canals open, and closes by ~P15[12, 13]. Thus, the closure of the auditory CP to pure tones occurs before adolescence begins[14, 15]. Yet, while some studies have found that simple neuronal properties in the ACx (like response properties to pure tones) are nearly mature by adolescence, others report continued maturation well beyond this developmental stage[16, 17]. For example, previous studies found that auditory learning and behavioral performance in auditory discrimination tasks involving amplitude modulation detection, as well as temporal interval discrimination tasks remain highly constrained during adolescence[4, 5]. In those tasks, ACx neurons in adolescents exhibit high neuronal variability and lower tone sensitivity as compared to adults[4].

The perspective shift from viewing development as involving a set of discrete CPs to a broader view of continuous development with multiple, overlapping, CPs that affect each other, calls for studies to test the extent of functional modulation in neuronal features after their CP 'closes'. Thus, although the CP for pure tone representation in the ACx ends by ~P15, we hypothesized that learning of pure tone discrimination would remain malleable in adolescence, because other perceptual features and particularly other cognitive features are still developing. Further, it remains unclear how these features are expressed when measured in the context of active behavior. We measured cortical activity in the ACx of adolescent mice engaged in an auditory discrimination task involving pure tones. Our findings show that learning, behavior, and sound representations to pure tones in the ACx are not fully mature during adolescence.

Results

Adult mice outperform adolescents on a difficult, but not an easy, auditory discrimination task of pure tones

To study learning and perceptual performance abilities we trained and tested adolescent and adult mice on an auditory pure-tone discrimination Go/No-Go paradigm. For this purpose, we utilized an automated behavioral platform called the ‘Educage’ (Fig. 1a) [18]. We trained adolescent ($n = 15$) and adult ($n = 15$) mice from post-natal day 20 (P20)-P37 and P60-P77, respectively (Fig. 1b). Following a tone association period (3-days; exposure to Go tones when mice enter the drinking port), mice learned to discriminate between two pure tones separated by 1 octave (7.07kHz vs 14.14kHz; Fig. 1b-c, light blue; herein referred to as the ‘easy task’). After one week of training on the easy task (P23 to P30 in pre-adolescent mice and P63 to P70 in adult mice) mice were introduced to a second pure-tone pair, now separated by 0.25 octave (9.2kHz vs 10.9kHz; Fig. 1b-c, dark blue; herein referred to as the ‘hard task’). We completed the experiment after 1 week of learning on the easy and hard task simultaneously (i.e., at P37 for adolescents and P77 for adults). During the whole period of training, we also presented mice with low-probability, unrewarded ‘probe’ trials of pure-tones spanning the range of learned stimuli (Fig. 1c, grey). Figure 1d shows example learning curves (of d') from one adolescent and one adult mouse over the 14-day course of tone discrimination. Adolescents and adults learned the procedure of the task similarly well, as measured by the number of trials it took mice to reach a discriminability threshold of d' value ≥ 1 (Fig. 1e).

We tested for differences in auditory discrimination by comparing the performance of mice at different epochs along training. First, by the end of the first week of training, all mice performed above the discrimination threshold (Fig. 1f). Although adolescent mice performed significantly fewer trials during the first week of training (Fig. 1f; vertical lines: mean # of trials), there was no significant difference between the performance of adult and adolescent mice at the end of the first week (Fig. 1f; horizontal lines: mean d' ; d' calculated from the last 100 trials). Following one week of training on the easy task, we introduced the second pair of pure tones (Fig. 1b; ‘Easy + Hard’). Immediately after introducing the hard pair, the performance of adolescent mice on the easy task dropped, while those of adults remained unaffected (Fig. 1g). During the second week of training, adolescents regained their high d' values on the easy task, such that by the end of the experiment both groups had similarly high performance on the easy task (Fig. 1h).

We then compared how mice learned the hard task during the second week while training on both the easy and hard task simultaneously. While both groups now performed a similar number of trials (Fig. 1i, vertical lines: mean # of trials), adult mice outperformed adolescent mice on the hard task (Fig. 1i, horizontal lines: mean d'), suggesting that adolescents struggled with this harder level of discrimination. Indeed, on average, adolescent mice did not improve on the hard task during the second week of training (Fig. 1j, ‘adolescent’). Adults, on the other hand, improved their performance on the hard task during the second week (Fig. 1j, ‘adult’). Notably, neither group could learn the hard task if it was not preceded by the easy task (Fig. S1-1a, b).

Each mouse performed a unique number of trials, as these were initiated by the mice spontaneously. As noted above, pre-adolescent mice performed fewer trials/day as compared to adults but, notably, only during the first week of training (mean $\pm$ STE: adolescents, 8317 ± 712 trials, adults 13229 ± 514 trials, t -stat = -5.6314 , $df = 28$, $p = 4.9566e-06$, two-sample independent t -test). Still, this difference raises the question whether the different number of trials affected our conclusions from the first week of training. Thus, to evaluate the possible connection between the number of trials and performance in our data, we carried out several analyses. First, we found

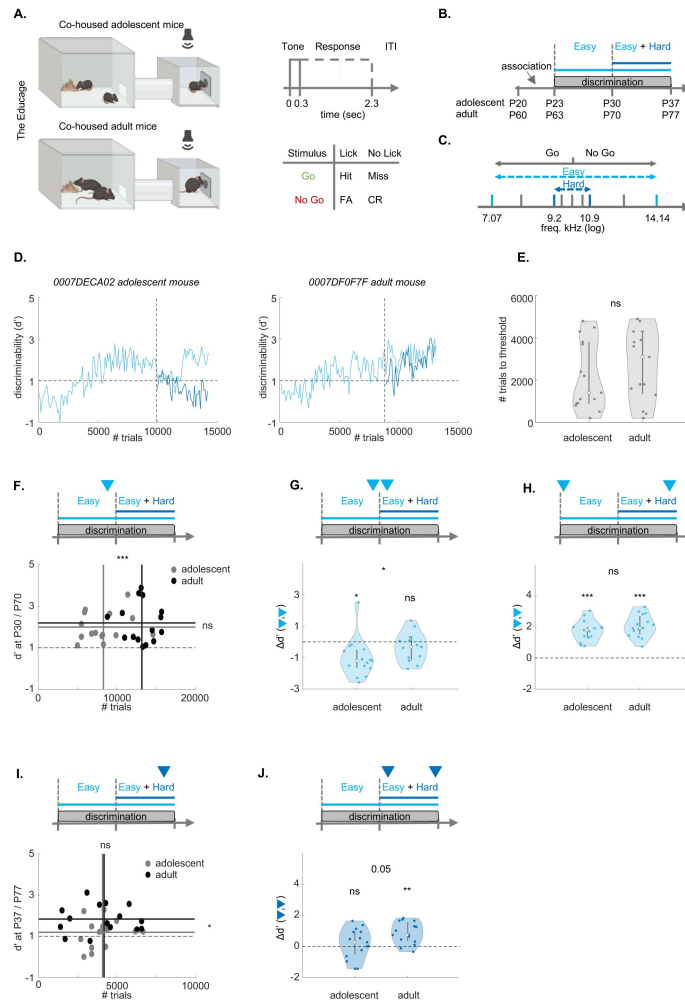


Figure 1

Adolescent mice exhibit lower performance during self-initiated auditory learning in the 'Educage'.

A. Schematic model of the 'Educage' (left), the trial structure and trial types (FA, false alarm; CR, correct reject). Created with *BioRender.com*. **B.** Experimental timeline. Total training time was 21 days (± 2). **C.** The sounds used for training. Light blue: easy task; Dark blue: hard task; Grey: catch-trials. **D.** Learning curve examples. Adolescent mouse (left, P20-to-P37); adult mouse (right, P60-to-P77). Vertical dashed lines indicate the easy-hard transition. Horizontal line is $d' = 1$. **E.** Number of trials to reach threshold ($d' \geq 1$; adolescents, $n = 15$; adults, $n = 15$; $z = -0.1659$; $p = 0.8682$, two-sample Wilcoxon rank sum test). **F.** Discriminability (d') of the easy task in adolescent mice at P30 (grey; $n = 15$, $d' = 2.0001 \pm 0.1791$; # trials = 8317 $\pm$ 712) and adult mice at P70 (black; $n = 15$, $d' = 2.1986 \pm 0.2441$; # trials = 13229 $\pm$ 514, top: marked by the arrowhead). Dashed lines: mean trials per group (t -stat = -5.6314, $df = 28$, $p = 4.9566e-06$, two-sample independent t -test, solid vertical line), and mean d' per age group ($z = 0.2074$; $p = 0.8357$, two-sample Wilcoxon rank sum test, solid horizontal line). **G.** Change in discriminability ($\Delta d'$) of the easy task before and after the introduction of the hard task (Top: arrowheads; left: adolescent, signed rank = 14, $p = 0.0067$; right: adult, signed rank = 14, $p = 0.1514$, one-sample Wilcoxon signed rank test; $\Delta d'$ between adult and adolescent mice: $z = -2.0739$; $p = 0.0381$, two-sample Wilcoxon rank sum test). **H.** Same as 'G' for the first 100 and last 100 trials of the experiment in the easy task (adolescent signed rank = 120, $p = 6.1035e-05$; adult signed rank = 120, $p = 6.1035e-05$, one-sample Wilcoxon signed rank test; $\Delta d'$ between adult and adolescent mice: $z = -1.0370$; $p = 0.2998$, two-sample Wilcoxon rank sum test). **I.** Same as 'F' for the hard task (adolescent-grey; $n = 15$, $d' = 1.1895 \pm 0.1783$; # trials = 4163 $\pm$ 297; adult-black; $n = 15$, $d' = 1.8342 \pm 0.1743$; # trials = 4102 $\pm$ 475; mean trials per group: t -stat = 0.1306, $df = 28$, $p = 0.8970$, two-sample independent t -test, solid vertical line; mean d' per group: $z = -2.2398$; $p = 0.0251$, two-sample Wilcoxon rank sum test, solid horizontal line). **J.** Same as 'G' for the hard task. (adolescent signed rank = 73, $p = 0.4887$; adult signed rank = 114, $p = 8.5449e-04$, one-sample Wilcoxon signed rank test; $\Delta d'$ between adult and adolescent mice: $z = -1.9495$; $p = 0.0512$, two-sample Wilcoxon rank sum test).

that the number of trials and d' were not correlated in either age-group during the first week of discrimination on the easy task (Pearson- r ; adolescent: $r = 0.1880$, $p = 0.5023$; adult: $r = -0.0730$, $p = 0.7959$), nor on the second week of discrimination on both the easy and the hard task (Pearson- r ; adolescent: $r = 0.2264$, $p = 0.4171$; adult: $r = 0.0194$, $p = 0.9454$). Second, we compared performance at three time points along the task with reference to times with shared number of trials: 1) evaluating d' after the minimal number of trials performed by all mice during the first week of training, and 2) evaluating d' at the mean number of trials of each group as a reference. We found no significant differences between adolescents and adults in all comparisons (**Fig. S1-1c-e**). Additionally, we found no significant differences between males and females, nor among experiments that had different numbers of co-housed mice in the Educage (**Table 1**). Taken together, despite the different absolute number of trials during the first week, we conclude that adolescents and adults learned the easy task equally quickly and effectively. The central difference between the age groups was in their ability to learn the hard task during the second week, when adults outperformed adolescents.

Adolescents have higher lick bias and higher behavioral variability

We next asked what perceptual and/or cognitive aspect of behavior is different among the age groups. We plotted psychometric curves based on both learned and probe trials from the last 2000 trials of the experiment (**Fig. 2a**). The false alarm rate but not the hit rate of adolescent mice was significantly higher as compared to adults (**Fig. 2b**). To calculate the decision boundary of mice and their perceptual sensitivity we normalized the psychometric curves of each mouse using a unity-based normalization and sigmoid fitting (**Fig. 2c**). We found no differences in the decision boundary or slopes of the curves, suggesting that perceptual sensitivity is not different between the groups (**Fig. 2c**). To test whether the differences arise from a cognitive effect (e.g., lick bias), we calculated the decision threshold as the maximal lick bias, also known as the criterion bias (C-bias). C-bias reflects the tendency to respond in a liberal (i.e., negative C-bias) or a conservative (i.e., positive C-bias) manner to the sounds during the task. C-bias was negative for both groups, yet significantly lower in adolescents, suggesting that their cognitive ability to withhold licking is inferior to that of adults (**Fig. 2d**).

Given that adolescents and adults show different dynamics in their performance along the task (**Fig. 1**), we tested whether the differences we found in the lick bias are consistent at different times along the task. To normalize for different numbers of trials in each mouse, we divided their training episodes into the first and last tertiles during the first week and second week of training separately (**Fig. 2e**). The maximal lick bias was significantly higher in adolescents in all but the very first training episode, when mice learned the procedure (**Fig. 2f**). These data highlight the general cognitive difficulty of adolescents to withhold licking during the task.

The cognitive sensitivity of adolescents was further evident when their performance dropped right after the task was switched from 'easy only' to 'easy + hard' (**Fig. 1f**). The lick bias of adolescent mice became significantly stronger after the introduction of the hard task (**Fig. 2g**). In contrast, the lick bias of adult mice remained stable (**Fig. 2g**). These data further suggest that this cognitive trait (i.e., to withhold licking) has not yet matured in the young mice. To test if the lick bias was stimulus-related or a result of general impulsivity, we also compared the inter-trial interval (ITI) after different trial outcomes, and observed similar ITIs across groups, except for false alarm trials, where adolescents initiated the trials significantly faster (**Fig. S2-1**). This, too, suggests that adolescent mice exhibit more impulsive responses following punishments (or alternatively are less impacted by punishments), contributing to their immature lick bias.

Prior studies have shown that a dominant feature of adolescent behavior is that it is more variable compared to adults[4]. To examine behavioral variability in our data, we calculated the coefficient of variation (CV) of the discriminability (d') for each mouse across training. While there was no significant difference in d' variability when training on the easy task during the first week

Fixed effects	Estimate	STE	T-statistic	DF	P-value	CI (lower)	CI (upper)
Intercept	1.6001	0.2312	6.9222	27	3.8798e-07	1.1258	2.0744
Age	0.5994	0.2266	2.6457	27	0.0269	0.1346	1.0643
Sex	-0.0428	0.2476	-0.1729	27	0.9999	-0.5509	0.4653

Table 1

Behavioral differences between adolescent and adult mice are age-, but not sex-related.

Fixed effects of age and sex, and the random effects of co-housing in the 'Educage' on the discriminability (mean d' of the last 100 trials of the easy and hard task to avoid pseudo replication) of all mice (Number of observations = 30, Fixed effects coefficients = 3, Random effects coefficients = 7, Covariance parameters = 2). Coefficient estimates, STE, T-statistic, degrees of freedom, p-values (adjusted for multiple comparisons with the Bonferroni method) and the lower and upper Confidence Interval (95%). The model includes random effects coefficients of the Cage ID in each group of co-housed mice (7 cages in total; see *methods*, equation 7). Model structure: discriminability(d') $\sim$ age + sex + (1 | cage ID).

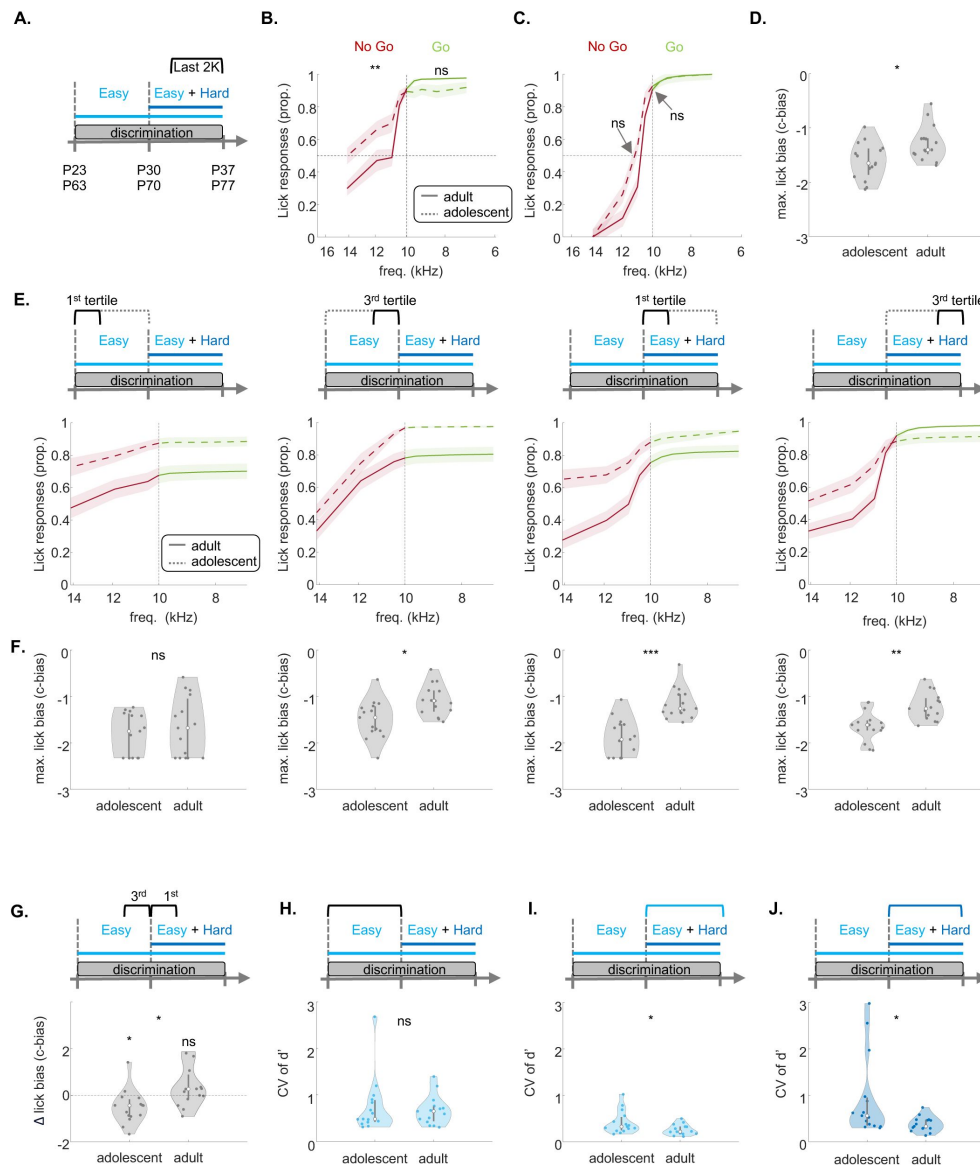


Figure 2

Adolescent mice exhibit lower performance in the head-fixed discrimination task.

A. Experimental timeline of training followed by recordings. **B.** Trial structure during the recording. Solid lines indicate the tone period. Dashed lines show the reward or punishment delay (0.6 sec), and the response window (2 sec). **C.** Example session. Licks (grey ticks) and trial outcomes (hit = green, false alarm = yellow, miss = red and correct reject = blue) across all trials in one recording session. **D.** Discriminability during training sessions for the easy task (light blue) and hard task (dark blue). **E.** Change in d' after the introduction of the hard task (last 100 trials of the last session of the easy task compared to last 100 trials of the first hard session; adolescent: sign-rank = 1, $p = 0.0625$; adult: sign-rank = 10, $p = 0.9998$; rank-sum = 14, $p = 0.0381$, two-sample Wilcoxon rank sum test). **F.** Expert d' of the last 100 trials during the last training session of the easy task (rank-sum = 21, $p = 0.1255$, two-sample Wilcoxon rank sum test). **G.** Same as 'F', but for the hard task (rank-sum = 17, $p = 0.0173$, two-sample Wilcoxon rank sum test). **H.** Behavioral performance (average d' of the easy and the hard task) per mouse during recording sessions for adolescents ($n = 13$, left) and adults ($n = 14$, right; trials per recording: adolescent: 340.5385 ± 45.0650 ; adult: 431.1429 ± 30.3367 ; independent t-test, t-statistic = -203.7581, $p = 0.1116$). **I.** Same as 'H' but only for the first 148 trials. The color bar shows the p-values between the groups. **J.** Average cumulative licks per trial in adolescents (dashed-line) and adults (solid-line) from -200ms before tone-onset until the reward or punishment delay, 500ms after tone-offset. **K.** Lick latency per trial for adolescent (left) and adult (right) groups during electrophysiological recordings (LME statistics are shown in [supplemental Table 1](#)). **J.** Same as 'K' for the Lick count.

of training (**Fig. 2h**), adult mice showed significantly lower variability during the second week of training while training on the ‘easy + hard’ versions of the task (**Fig. 2i,j**). Thus, adolescent learning and behavior are characterized by both lower response inhibition and higher variability, particularly during the second week when the task involved more challenging discrimination for the younger mice.

Adult mice outperform adolescent mice on a head-fixed discrimination task

To enable electrophysiological recordings during expert task performance, we trained adolescent ($n = 5$) and adult ($n = 6$) mice on a head-fixed learning paradigm, using a similar protocol as in the Educage (**Fig. 3a**). In the head-fixed protocol, water access was limited to training sessions. Head-fixed mice were trained to lick after a 100ms tone and were rewarded (with water) or punished (by a 2 sec white noise) only after a 600ms delay (**Fig. 3b**). Each session was concluded after the mice became satiated and stopped licking. Water supply was limited to 0.0125 ml per day per gram of body weight, and mice that did not consume this amount were compensated after training (see Methods; **Fig. 3c** shows an example session from one mouse). The number of trials per training session was significantly lower in adolescents (mean $\pm$ STE: adolescents: 410 ± 35 , adults: 580 ± 37 t-statistic = -27.9855 , $p = 0.0012$, independent two-sample t-test). However, the number of sessions to reach the behavioral threshold ($d' \geq 1$ in the easy task) was not significantly different between adolescents and adults (mean $\pm$ STE: adolescents: 3.7 ± 0.8 , adults: 3.8 ± 0.7 ; t-statistic = -0.8531 , $p = 0.1971$, independent two-sample t-test, **Fig. 3d** shows all learning curves). Similar to learning in the Educage, adolescent performance (but not that of adults) decreased after we introduced the hard task (**Fig. 3e**). Another similarity between the head-fixed and Educage versions of the task was that we found no significant differences between adolescents and adults in the performance of the easy task (**Fig. 3f**), and that adults outperformed adolescents on the hard discrimination (**Fig. 3g**). Also similar to the behavior in the Educage, lick responses of adolescent mice were generally higher (**Fig. S3-1a**), and adolescents exhibited a stronger lick bias (**Fig. S3-1b**). In line with this result, in the head-fixed task we found higher impulsiveness in the proportion and the number of licks during the inter-trial intervals after FA trials in adolescents (**Fig. S3-1c, d**).

To test whether ACx was necessary for expert discrimination in this task, we performed a causal experiment in adult mice. Adult mice ($n=3$) were injected bilaterally with AAV5-CAMKII-GtACR2-FRED_kv_2.1 into the ACx, and optical fibers were implanted over the injection sites (**Fig. S3-2a, c**; **Fig. S3-3**). This viral construct enables optogenetic silencing of excitatory neurons via GtACR2 (*Guillardia theta* Anion Channelrhodopsin 2), under the control of the CaMKII promoter. It includes a red fluorescent protein for expression verification (FRED) and a soma-targeting motif derived from the Kv2.1 potassium channel to enhance membrane localization.

Mice were trained identically on the head-fixed Go/No-Go paradigm as described above. At the end of training, we attached a light source to the implanted fibers to allow optogenetic suppression during several sessions (**Fig. S3-2a**). In these sessions, light was applied in 50% of trials in a pseudo-random fashion, with light-on trials starting from -50 ms prior to tone onset up to 50 ms after tone offset (**Fig. S3-2b**). A control group of mice ($n= 3$) were injected with AAV9-CaMKII-dTomato and went through the exact same procedure (**Fig. S3-2a, c**). We compared the lick responses for both Go and No-Go stimuli under light-on and light-off conditions. We found that lick responses were strongly affected in the GtACR2-injected mice ($n = 13$ sessions), but not in controls ($n = 8$ sessions) (**Fig. S3-2d**). Light-on conditions only affected the lick responses of the inhibited trial and not the light-off trials (**Fig. S3-2e**). Discriminability (d') decreased significantly with optogenetic suppression for both easy and hard tasks, although all mice still performed better in the easy task under both conditions (**Fig. S3-2f**). These results confirm findings by others in

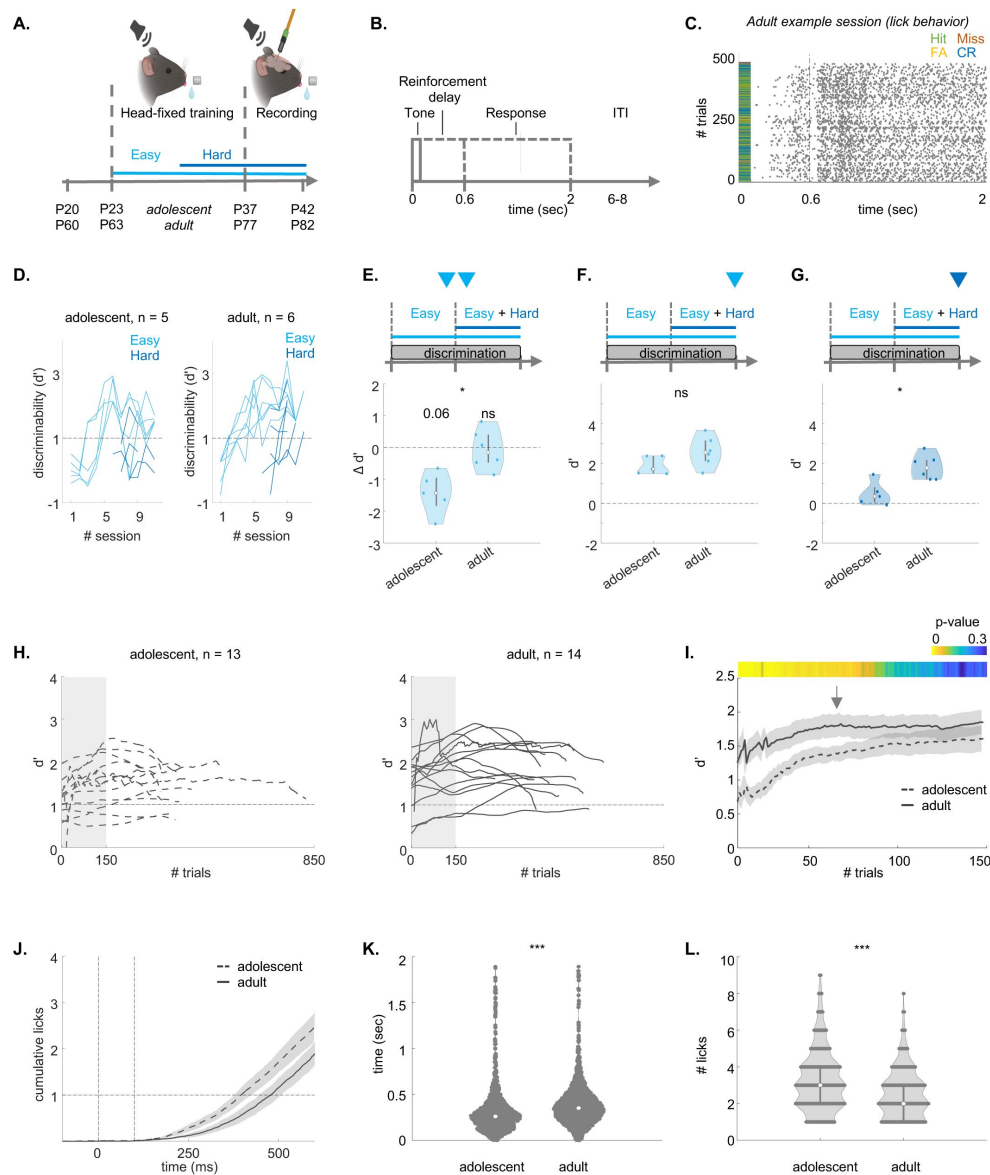


Figure 3

Adolescent mice exhibit lower performance in the head-fixed discrimination task.

A. Experimental timeline of training followed by recordings. Created with *BioRender.com*. **B.** Trial structure during the recording. Solid lines indicate the tone period. Dashed lines show the reward or punishment delay (0.6 sec), and the response window (2 sec). **C.** Example session. Licks (grey ticks) and trial outcomes (hit = green, false alarm = yellow, miss = red and correct reject = blue) across all trials in one recording session. **D.** Discriminability during training sessions for the easy task (light blue) and hard task (dark blue). **E.** Change in d' after the introduction of the hard task (last 100 trials of the last session of the easy task compared to last 100 trials of the first hard session; adolescent: sign-rank = 1, $p = 0.0625$; adult: sign-rank = 10, $p = 0.9998$; rank-sum = 14; $p = 0.0381$, two-sample Wilcoxon rank sum test). **F.** Expert d' of the last 100 trials during the last training session of the easy task (rank-sum = 21; $p = 0.1255$, two-sample Wilcoxon rank sum test). **G.** Same as 'F', but for the hard task (rank-sum = 17; $p = 0.0173$, two-sample Wilcoxon rank sum test). **H.** Behavioral performance (average d' of the easy and the hard task) per mouse during recording sessions for adolescents ($n=13$, left) and adults ($n=14$, right; trials per recording: adolescent: 340.5385 ± 45.0650 ; adult: 431.1429 ± 30.3367 ; independent t-test, t-statistic = -203.7581 , $p = 0.1116$). **I.** Same as 'H' but only for the first 148 trials. The color bar shows the p-values between the groups. **J.** Average cumulative licks per trial in adolescents (dashed-line) and adults (solid-line) from -200 ms before tone-onset until the reward or punishment delay, 500 ms after tone-offset. **K.** Lick latency per trial for adolescent (left) and adult (right) groups during electrophysiological recordings (LME statistics are shown in **supplemental Table 1**). **L.** Same as 'K' for the Lick count.

similar (though not identical) tasks [19, 20], suggesting that the ACx is necessary for executing the behavior during the expert stage of this task (but see Discussion for limitations of this experiment).

To measure neural responses during behavioral performance, head-fixed expert mice underwent multiple recording sessions targeting the ACx (adults, $n = 14$ recording sessions from 6 mice, age: P77-P82; adolescents, $n = 13$ recordings from 5 mice, age: P37-P42). During the recordings, the performance of all mice on the easy task was at least $d' > 1$, and the d' on the hard task was smaller for both groups (often smaller than 1; **Fig. S3-4a**). Performance was not significantly different across the multiple recording sessions in each mouse (**Fig. S3-4b**). Behavioral performance during recordings was not significantly different to performance in the Educage (**Fig. S3-4c**). **Figure 3h** shows the evolution of behavioral discriminability (d') during all recordings for adolescents and adults. The performance was heterogeneous across mice as well as within recording-sessions. However, the total number of trials performed was not significantly different between age groups (trials per recording: adolescent: 340.5 ± 45 ; adult: 431.1 ± 30.3 ; independent t-test, t -statistic = -203.7581 , $p = 0.1116$).

We found additional differences between the behavior of adults and adolescents in the head restrained configuration. The average adolescent discriminability was lower at the beginning of the session (initial 78 trials) but then rapidly leveled out (**Fig. 3i**). **Figure 3j** shows the cumulative lick curves of adults and adolescents. Adolescent mice had shorter lick latencies (**Fig. 3k**), and higher lick counts (**Fig. 3l**; **Supplementary Table 1**). While the lick latencies were independent of the discriminability (d'), we found a significant interaction effect between lick count and the d' (**Supplementary Table 1**). Together, these behavioral differences suggest that the weaker response inhibition (here, expressed as lick latency) and higher reward anticipation (here, expressed as lick count) may contribute to the weaker performance of adolescents during the beginning of the recording session.

Neuronal representations of stimulus- and choice-related activity are immature in adolescents

To compare the activity of neurons in adolescent and adult mice, we recorded spiking activity from the ACx using the high-density Neuropixels-1 probes in expert mice engaged in the task. We targeted the ACx by inserting the probe in a diagonal angle, traversing four auditory regions: Dorsal Auditory Cortex (AUDd), Primary Auditory Cortex (AUDp), ventral auditory cortex (AUDv), and Temporal Association Cortex (Tea; **Fig. 4a**; **Fig. S4-1a,b**, show an example voltage trace and one example recording). We recorded multiple times from each mouse and used DiI- or DiO-coatings for different penetrations to verify probe trajectories postmortem. We used the 3D-allen CCF-slice reconstruction platform (github.com/cortex-lab/allenCCF), for high resolution anatomical registration (**Fig. 4b**; **Fig. S4-1c, d**).

We limited our single neuron analysis to well-isolated single units from infragranular layer 5 and layer 6 (recordings were largely restricted to these layers due to the probe angle), as two main projection layers of the cortex, which are key nodes for assessing cortical outputs [21]. Recorded neurons from other layers and areas were excluded from the analysis. Together, we collected data from 1145 single neurons in 13 recordings of 5 adolescent mice and 1267 single neurons in 14 recordings of 6 adult mice (**Supplementary Table 2**; an overview of every dataset per figure is detailed in **Supplementary Table 9**). We further restricted our analysis to single neurons excited by sounds during the first 150ms after tone onset. **Figure 4c** shows the normalized firing rate (FR) compared to the normalized lick rate (LR) from -200 ms to 600 ms after the tone onset, before the end of the reinforcement delay (**Fig. S4-1e** shows the FR in Hz per auditory sub-region). FR-PSTHs and LR-PSTHs were not correlated during the first 150ms (adolescents $r = -0.1515$, $p = 0.9878$; adults $r = -0.0638$, $p = 0.9942$). In total, we analyzed $n = 463$ neurons from adolescent mice and $n = 599$ from adults.

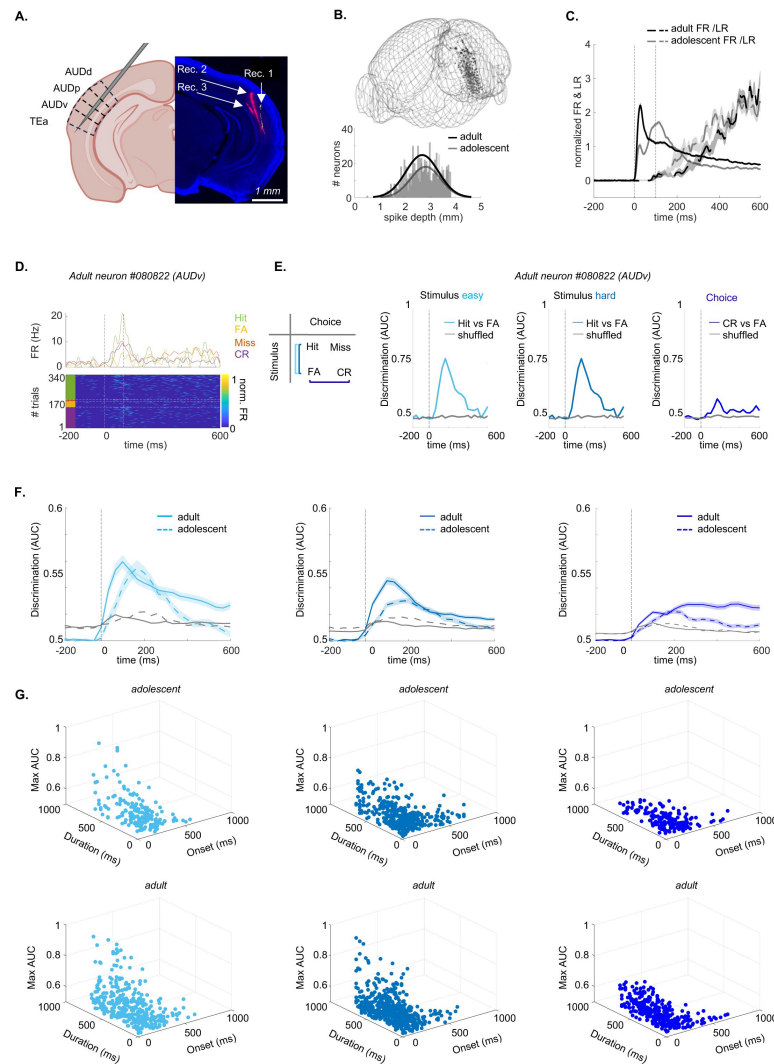


Figure 4

ACx neurons in adolescents exhibit lower discriminability in stimulus- and choice-related activity.

A. Recordings in ACx when the mouse is engaged in the task, using Neuropixels-1 probes. Left: Recordings were performed in AUDd, AUDp, AUDv, and TEa. Right: Fluorescent micrograph of a coronal brain slice showing the probe tracks of three recordings (red = DiI, yellow = DiO). Created with *BioRender.com*. **B.** Top: 3D-Reconstruction of recording sites in adolescents ($n = 13$; grey) and adults ($n = 14$; black). Bottom: distribution of the spike-depth of all excitatory tone-responsive L5/6 neurons in adolescents ($n = 455$; grey) and adults ($n = 607$; black). **C.** Normalized PSTH (FR in Hz) and lick-rate (LR in Hz) from -200 ms to $+600$ ms after tone-onset in adolescents (grey) and adults (black). **D.** Spiking activity from one example neuron sorted by trial outcome (hit, miss, false alarm, correct reject). Top: PSTH per trial outcome. Bottom: Heat map of the FR sorted per trial outcome. **E.** Discriminability values (AUC) over time (from -200 ms to 600 ms after tone onset) for one example neuron (same neuron as in 'D'). AUC values are shown for stimulus related activity (left: easy task, middle: hard task) and choice-related activity (right). Shuffled distribution in all curves is shown in grey. **F.** Same as 'E' for all neurons. The curves are average ($\pm$ STE) neuronal discriminability of adult neurons (solid line) and adolescent neurons (dashed line), for easy (adolescent neurons = 190, mice = 4, recordings = 7; adult $n = 358$, mice = 4, recordings = 8; left) and hard stimulus-related activity (adolescent $n = 429$; adult $n = 562$, mice = 5, recordings = 9; middle), and choice-related activity (adolescent $n = 429$; adult $n = 562$, mice = 5, recordings = 9; right). **G.** 3D plots of the onset-latency of discriminability (ms), duration of discriminability (ms), and maximal discriminability (AUC) of all neurons that showed significant discriminability. Left: easy task (adolescent neurons = 178 (93%), mice = 4, recordings = 6; adult $n = 346$ (97%), mice = 4, recordings = 8; left); Center: hard task (adolescent neurons = 399 (93%), mice = 5, recordings = 10; adult $n = 544$ (97%), mice = 6, recordings = 12; middle); Right: choice-related activity (adolescent neurons = 181 (95%), mice = 4, recordings = 9; adult $n = 339$ (95%), mice = 4, recordings = 7; right).

In **Supplementary Table 2** [↗](#), we present the number of neurons in our dataset, categorized by region. A detailed overview of the differences in firing properties per auditory region is also provided in **Supplementary Table 3** [↗](#) and **Supplementary Table 4** [↗](#).

The differences in firing properties between adolescent and adult neurons varied across auditory subregions. AUDp showed differences in all firing properties that we analyzed. In AUDd, adolescent neurons had greater latency to peak, full-width half-maximum, minimal latency, and lower trial-responsiveness. AUDv also showed greater minimal latency and lower lifetime sparseness. TEa neurons of adolescents exhibited greater minimal latency and higher spontaneous FR. While these differences may partially reflect the varying number of neurons recorded per subregion and different functional attributes of each region, we observed similar trends in Cohen's D values (effect sizes) across all four areas, supporting a generally consistent group differences in activity patterns (**Supplementary Table 3** [↗](#)). Consistent with previous work, we also found that firing properties such as minimal latency differed between auditory regions in both age groups (**Supplementary Table 4** [↗](#)) [22].

To study how adolescents and adults encode task performance we divided neuronal responses by trial outcome (**Fig. 4d** [↗](#) shows one example). Specifically, we assessed neuronal discriminability between stimulus-related and choice-related activity by calculating the area under the curve (AUC) from a receiver operating curve (ROC). Specifically, stimulus-related activity was calculated as the difference between hit and false alarm trials in the easy and the hard task separately (**Fig. 4e** [↗](#), left). Choice-related activity was calculated as the average difference between false-alarm and correct reject trials (**Fig. 4e** [↗](#), left). Miss trials were excluded since we had insufficient trial numbers of this outcome. **Figure 4e** [↗](#) (right) shows the average AUC values over time from an example neuron (same neuron shown in **Fig. 4d** [↗](#)), and **figure 4f** [↗](#) shows the average calculated from all neurons. The vast majority of neurons successfully discriminated (AUC above shuffled data) either or both stimulus- and choice-related activity (adolescent: easy stimuli = 93%, hard stimuli = 93%, choice = 95%; adult: easy stimuli = 97%, hard stimuli = 97%, choice = 95%). The onset-latency of discriminability was significantly slower, maximal discriminability significantly weaker, and the duration of discriminability significantly shorter in adolescent neurons compared to adult neurons (discriminability traces are plotted in **Fig. 4f** [↗](#), and all values from individual neurons and the 3 metrics are plotted together in **Fig. 4g** [↗](#); see **Table 2** [↗](#) for statistics). The four auditory subregions were not significantly different in all 3 measured metrics (**Fig. S4-2** [↗](#)). Thus, auditory processing of stimulus- and choice-related activity in the adolescent ACx has not yet reached full maturity.

Adults exhibit better population decoding of the difficult task compared to adolescents

While individual neurons exhibited weaker discriminability in adolescents than in adults, at a population level these effects could saturate out. We therefore tested the ability of recorded populations to decode trial outcomes. We used a linear discriminant analysis (LDA) decoder to quantify hit versus correct reject trial outcomes. We applied stringent inclusion criteria for the number of simultaneously recorded neurons (see Methods), which restricted our analysis to hit versus correct rejection trials with sufficient trial counts for statistical reliability. We note that this comparison extracts information of the combination of stimulus and choice information indiscriminately (i.e. collectively it includes the full 'task-related' differences). We decoded separately in easy trials and hard trials, based on population activity during the first 200 milliseconds. For the easy task there was no significant difference between decoding accuracy in adolescents and adults (**Fig. 5a** [↗](#)). However, decoding accuracy was significantly higher in adults compared to adolescents in hard trials (**Fig. 5a** [↗](#)). For both adults and adolescents, decoding accuracy was significantly lower in hard trials compared to easy trials (**Fig. 5a** [↗](#)). We performed finer timescale decoding to quantify the onset latency of differences between the activity for hit vs.

Stimulus easy							
Fixed Effects	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	163.6104	7.9457	20.591	518	6.29587E-69	148.0009	179.2199
Onset (ms)	-29.7978	9.7626	-3.0523	518	0.00716101	-48.9765	-10.6191
Intercept	216.2921	13.2858	16.2799	518	5.15393E-48	190.1919	242.3924
Duration (ms)	63.2599	16.3499	3.8691	518	0.000369146	31.1401	95.3796
Intercept	0.6035	0.007	85.6404	518	0.0001	0.5897	0.6174
Max AUC	0.0186	0.0049	3.7564	518	0.0001	0.0088	0.02841
Stimulus hard							
Fixed Effects	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	151.1905	5.4684	27.6479	535	3.6685E-123	140.4588	161.9222
Onset (ms)	-35.3357	7.1998	-4.9079	535	3.25392E-06	-49.4651	-21.2063
Intercept	220.614	8.4357	26.1525	535	2.8344E-113	204.0591	237.1689
Duration (ms)	55.6268	11.1065	5.0085	535	1.9647E-06	33.8305	77.4231
Intercept	0.5821	0.0034	170.2777	535	0.0001	0.5754	0.5888
Max AUC	0.02013	0.0048	4.1119	535	4.5413e-05	0.0105	0.0297
Choice							
Fixed Effects	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	174.5856	8.259	21.139	518	1.6732E-71	158.3605	190.8108
Onset (ms)	-24.3644	10.2288	-2.3819	518	0.052746707	-44.4595	-4.2693
Intercept	228.1768	12.92	17.6607	518	1.64628E-54	202.7947	253.5589
Duration (ms)	82.0002	16.0017	5.1245	518	1.26556E-06	50.564	113.4363
Intercept	0.5491	0.0027	200.1358	518	0.0001	0.5437	0.5545
Max AUC	0.01660	0.0034	4.8959	518	3.92843E-06	0.01	0.0233

Table 2

Neuronal discrimination is later, shorter, and less precise in adolescent neurons.

Linear mixed effect models of the neuronal discriminability in adolescence and adulthood per stimulus-related activity in the easy task (Number of observations = 524, Fixed effects coefficients = 2, Random effects coefficients = 10, Covariance parameters = 3), stimulus related activity in the hard task (Number of observations = 943, Fixed effects coefficients = 2, Random effects coefficients = 14), and choice-related activity (Number of observations = 520, Fixed effects coefficients = 2, Random effects coefficients = 10, Covariance parameters = 3). The table shows the fixed effects of the coefficient estimates, STE, T-statistic, degrees of freedom, p-values (corrected for multiple comparisons with Bonferroni-correction) and the upper and lower CI of the effect of age on the onset latency of discrimination, duration of discrimination and maximal neuronal discrimination (AUC). Each model also included random effect coefficients of each mouse, and recording per mouse. P-values for were adjusted with post-hoc tests using Bonferroni-correction (see *methods*, equation 9). Model structures: onset latency (ms) ~ age + (1 | Mouse ID) + (1 | Recording ID); duration (ms) ~ age + (1 | Mouse ID) + (1 | Recording ID); maximal discriminability (AUC) ~ age + (1 | Mouse ID) + (1 | Recording ID).

correct reject trials, defined as the time in which the decoding accuracy first crosses three times the standard deviation away from the baseline (see Methods). We found that the latency was shorter in adults in both the easy and hard trials (**Fig. 5b**). In adolescent animals the latency was larger in hard trials (**Fig. 5b**) but not in adult animals (**Fig. 5b**). We then extended this analysis over a temporal window from -0.5s to 10s relative to tone onset (**Fig. 5c**). In the 200ms window after the response period (2.0-2.2 s), there is no longer a significant difference in decoding accuracy between adult and adolescent mice as well as between the two task difficulties (**Fig. S5-1**).

To further quantify the different decoding performance between two tasks difficulties, we plotted the fisher separation metric ($\frac{\sigma_{between}^2}{\sigma_{within}^2}$, see Methods) for each level of difficulty. While decoding accuracy reflects how well a classifier can distinguish between trial types, the Fisher separation metric specifically quantifies the ratio of between-category variance to within-category variance. The separation ratio of hard over easy tasks was significantly higher in adult mice than in adolescent mice (robust linear regression, see Methods). Separation could have changed due to alterations in the mean or the dispersion around the mean. Testing the dispersion, we found that the variance around the means was not equal across easy and hard trials but rather significantly increased for the hard trials (**Fig. 5d-f**). However, there was no significant difference in variance between adolescent and adult mice within the same task (**Fig. 5d-f**).

Both age and learning affect cortical plasticity in mice engaged in the task

We next studied how learning contributes to cortical plasticity in the different age groups. To separate between age- and learning-related effects, we recorded from the ACx of two new groups of mice (adolescents and adults that are age-matched to expert mice) that are novice on the task (**Fig. 6a**, novice P37-P42, recordings= 6; novice P77-P82, recordings =6, n = 3 mice per group). We collected data from 130 tone-responsive (by excitation) neurons in novice adolescents and 186 tone-responsive (by excitation) neurons in novice adults (**Fig. 6b**; **Supplementary Table 2**). **Figure 6c** shows the normalized average FR PSTH together with the LR PSTH for novice mice. Notably, novice mice did not discriminate between sounds (**Fig. S6-1a, b**). Similar to expert mice (**Supplementary Table 3**), we found diverse firing properties across auditory subregions in novice mice (**Supplementary Table 5**). Also similar to expert mice, several basic firing properties were different across regions in both novice adolescents and novice adults (**Supplementary Table 6**).

To reveal learning-related differences, we compared neuronal responses in novice versus expert mice in both age-groups. We computed single-neuron AUCs for Go and No-Go trials (here we did not distinguish between stimulus and choice and computed AUCs for all trial outcomes, since expert mice mostly performed CR and novice mice mostly performed FA to No-Go tones), in both easy and hard task (**Fig. 6d-g**). The onset-latency for AUC discrimination was significantly different between adolescents and adults (shorter in adults), but not between novice and expert mice, nor between the easy and the hard task (**Table 3**). The maximal AUC discriminability was higher in adults compared to adolescents. It was also higher in experts than in novice, and in the easy compared to the hard task (**Table 3**). Similarly, the duration of discriminability was longer in adults and after learning (**Table 3**). There was a significant interaction between age and learning, as well learning and task difficulty in the maximal discriminability (**Table 3**). The onset-latency, duration of discriminability and maximal discriminability, were not significantly different among the AUDd, AUDp, AUDv and TEa in the four experimental groups (**Fig. S6-2**).

To compare neuronal discriminability and behavior more explicitly, we analyzed the correlations between neuronal decoding and behavioral performance using linear regression. In the easy task, both age groups showed high correlation between neuronal discriminability and behavioral

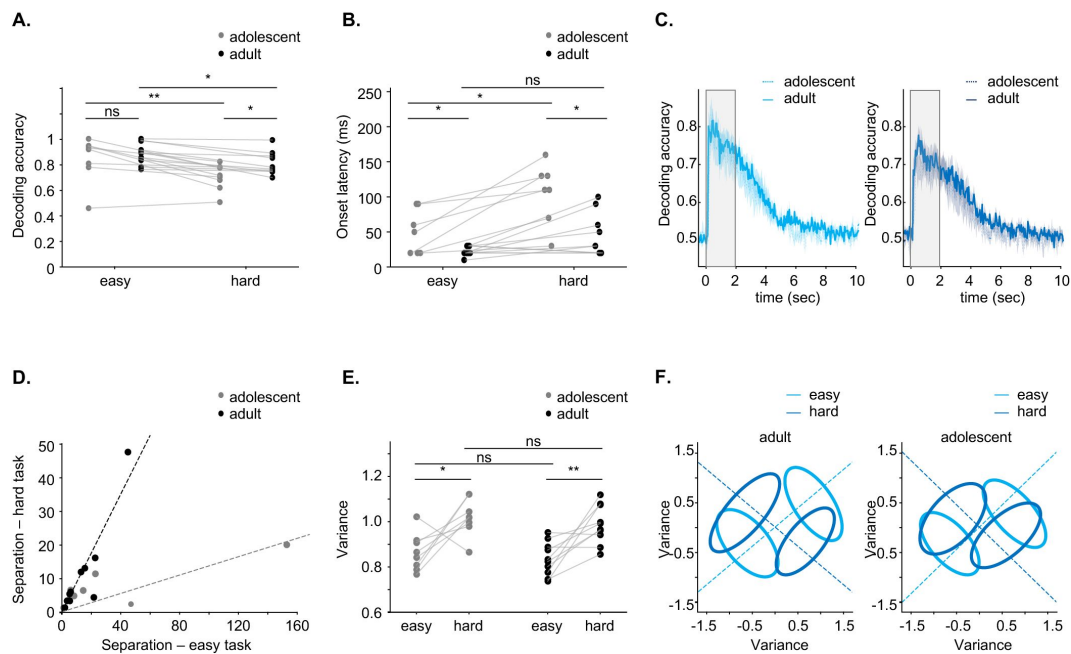


Figure 5

Decoding in adult neuronal populations outperforms decoding in adolescents.

A. Decoding accuracy for the first 200ms across all recordings in both adults (black) and adolescents (grey) for the easy task (adolescents compared to adults, $p = 0.5000$, Student's t-test) and the hard task (adolescents compared to adults, $p = 0.0300$, Student's t-test). Decoding is better in the easy task for both age groups (adults: $p = 0.0030$; adolescents: $p = 0.01$, paired t-test). **B.** Decoding latency for all recordings in the easy task ($p = 0.0200$, Student's t-test) and the hard task ($p = 0.0030$, Student's t-test), as well as compared between age groups (easy task, $p = 0.05400$, paired t-test; hard task: $p = 0.0100$). **C.** Decoding accuracy over a time window from -0.5 s to 10 s (the response window highlighted in the grey) for the easy task (left) and the hard task (right). **D.** LDA separation for easy and hard tasks. Lines represent robust linear regression fits without intercept (Huber loss; robust linear regression, $p = 0.0001$). **E.** Single trial variance for easy and hard tasks in adolescent and adult recordings (adults: $p = 0.0040$; adolescents: $p = 0.0300$, paired t-test; easy task: $p = 0.4500$; hard task: $p = 0.4100$, Student's t-test). **F.** Visualization of population representations for the stimuli in easy and hard tasks. Dotted lines indicate decoding dimensions, and ellipses represent the covariance of the representations.

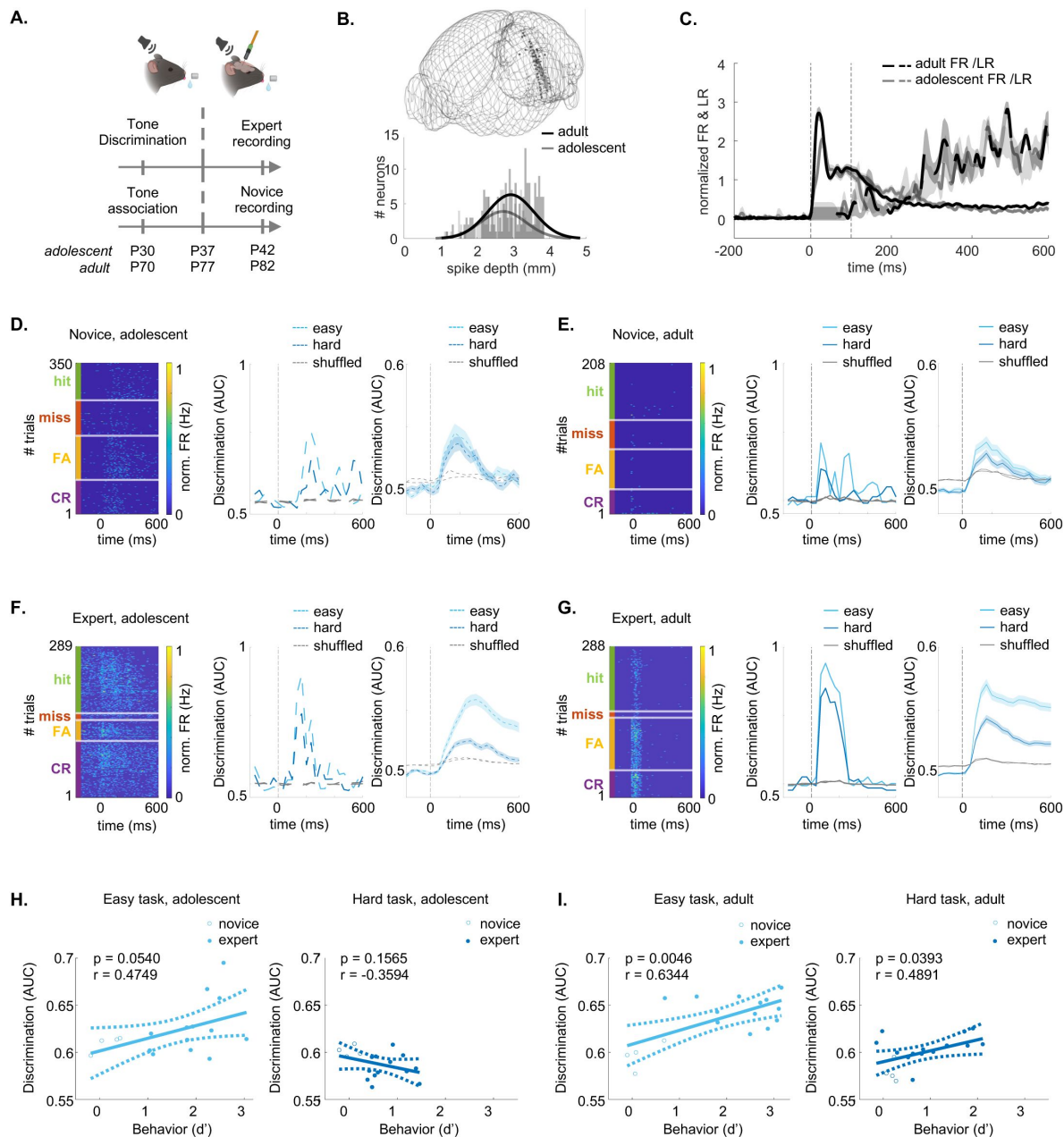


Figure 6

Cortical activity during behavior reflects both age- and learning-induced effects.

A. Training and recording schedule for novice mice, compared to expert mice. Created with [BioRender.com](https://www.biorender.com). **B.** 3D-Reconstruction of recording sites in novice adolescent (n = 6; grey) and novice adult (n = 6; black) mice. Bottom: spike-depth of excitatory tone-responsive L5/6 adolescent (n = 130; grey) and adult (n = 186; black) neurons. **C.** Normalized FR and lick rate (LR) PSTH from -200ms to 600ms after tone-onset in adolescents (grey) and adults (black). Average $\pm$ sem. **D.** Single neuron data from novice adolescent mice. Left: Heat map of the FR per trial from one example neuron sorted by trial outcomes. Center: the AUC of the neuron from the left for the easy and hard stimulus pairs (light and dark blue, respectively). Right: Average ($\pm$ -SEM) AUC of all neurons in the novice group (n = 140 neurons). **E-G.** Same as 'D' for novice adult (n = 186 neurons), expert adolescents (n = 455 neurons; Easy vs hard), and expert adults (n = 604 neurons; Easy vs hard). **H.** Linear regression analysis between the average AUC per recording and the behavioral d' during the recording (the correlation and p values are indicated for each plot). **I.** Same as 'I' for adult mice.

Fixed Effects Onset Latency(ms)	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	92.6519	3.8909	23.8125	2582	5.4024E-113	85.0223	100.2815
Age	-22.6378	5.0624	-4.4718	2582	2.42782E-05	-32.5645	-12.7111
Learning	5.6206	8.7235	0.6443	2582	0.9999	-11.4851	22.7263
Task difficulty	0.8813	5.4571	0.1615	2582	0.9999	-9.8194	11.5819
Age- Learning	-6.8831	11.1114	-0.6195	2582	0.9999	-28.6713	14.905
Age - Difficulty	0.2461	7.2185	0.0341	2582	0.9999	-13.9085	14.4006
Learning - Difficulty	-2.5145	12.3257	-0.204	2582	0.9999	-26.6837	21.6546
Age-Learning-Difficulty	4.4526	15.7691	0.2824	2582	0.9999	-26.4688	35.374
Fixed Effects Duration of Discrimination (ms)	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	299.7098	7.8305	38.2747	2582	8.4469E-254	284.3551	315.0645
Age	63.2136	9.2513	6.833	2582	3.1017E-11	45.0729	81.3544
Learning	-88.8699	15.9421	-5.5745	2582	8.21813E-08	-120.1306	-57.6092
Task difficulty	-53.9606	9.9745	-5.4098	2582	2.06623E-07	-73.5194	-34.4017
Age- Learning	-30.6998	20.3155	-1.5112	2582	0.3926	-70.5361	9.1365
Age - Difficulty	-18.8772	13.194	-1.4307	2582	0.4579	-44.749	6.9947
Learning - Difficulty	53.8178	22.5289	2.3888	2582	0.0509	9.6414	97.9943
Age-Learning-Difficulty	12.9698	28.823	0.45	2582	0.9999	-43.5488	69.4884
Fixed Effects Maximal Discrimination (AUC)	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	0.6278	0.0042	148.636	2582	0.0001	0.6195	0.6361
Age	0.024	0.0056	4.2953	2582	5.42797E-05	0.0131	0.035
Learning	-0.0249	0.0096	-2.5859	2582	0.029303232	-0.0438	-0.006
Task difficulty	-0.0438	0.006	-7.24	2582	1.76807E-12	-0.0556	-0.0319
Age- Learning	-0.03	0.0123	-2.4376	2582	0.0446	-0.0541	-0.0059
Age - Difficulty	-0.0023	0.008	-0.2846	2582	0.9999	-0.0179	0.0134
Learning - Difficulty	0.0394	0.0136	2.8893	2582	0.0116	0.0127	0.0662
Age-Learning-Difficulty	-0.0094	0.0175	-0.5387	2582	0.9999	-0.0437	0.0248

Table 3

The effect of age, learning and task difficulty on the latency, duration, and ability to discriminate tones in ACx neurons.

Linear mixed effect models of the effect of age, learning and task difficulty on onset-latency of discrimination, duration of discrimination and maximal discriminability (Number of observations = 2590, Fixed effects coefficients = 8, Random effects coefficients = 20, Covariance parameters = 3). The table shows the fixed effects of the coefficient estimates, STE, T-statistic, degrees of freedom, p-values (corrected for multiple comparisons with Bonferroni-correction) and the upper and lower CI. The model also includes random effects coefficients of each mouse (adolescent novice = 3, adult novice = 3, adolescent expert = 5, adult expert = 6) and recording per mouse (n = 3). P-values for were adjusted with post-hoc tests using Bonferroni-correction (see *methods*, [equation 1](#) [↗](#) 0). Model structures: onset latency (ms) ~ age* learning * difficulty + (1 | Mouse ID) + (1 | Recording ID); duration (ms) ~ age* learning * difficulty + (1 | Mouse ID) + (1 | Recording ID); maximal discriminability (AUC) ~ age* learning * difficulty + (1 | Mouse ID) + (1 | Recording ID).

discriminability (**Fig. 6h,i**; light blue). In the hard task, only adult neurons had significant correlations between neuronal discriminability and behavioral discriminability (**Fig. 6h,i**; dark blue).

Effects of age and learning on tuning properties in passively listening mice

Auditory learning has been shown to induce changes in the tuning properties of sounds in the ACx [18, 19, 21, 23–28]. To study whether learning-induced plasticity in ACx is distinct in adolescents and adults, we recorded from the same mice described above under passive listening conditions by simply extending the recording following the engaged session (**Fig. 7a**, **Supplementary Table 7**; Adolescents recordings = 4; Adults recordings = 4). These data were compared to data from a group of novice mice, recorded under passive listening conditions following their own engaged session (**Supplementary Table 7**; novice adolescent recordings = 6; novice adult recordings = 6).

We characterized the response profile of ACx neurons using a ‘frequency-response area’ protocol composed of twenty pure tones (4–40 kHz, spaced at 0.1661 octaves steps), each played for 100ms at five different attenuations (72–32 dB SPL). We limited our analysis to significant excitatory units only, determined as being auditory responsive in a window from tone onset to 50ms after tone offset at 62 dB SPL. We then focused our tuning analysis on responses to 62 dB SPL because this sound level was close to the intensity used during the behavioral task (72 dB SPL), while still avoiding potential ceiling effects observed at higher SPLs in recordings of mice not engaged in the task.

Our dataset included four groups. The two adolescent groups were: 1) adolescent novices, and 2) adolescent experts, with 92 and 80 neurons, respectively. The two adult groups were: 3) adult novices, and 4) adult experts, with 123 and 84 neurons, respectively. Neurons in all groups responded to pure tone as expected from previous studies in adult mice, including classical V-shaped frequency response areas [22, 29–32]. A representative example from a responsive neuron of an adolescent expert and one from an adult expert are shown in **Figure 7b,c**. To assess differences in pure-tone responses between auditory subregions, we quantified the firing properties of AUDp versus AUDv (**Supplementary Table 8**; TEa and AUDd were excluded since fewer than 20 units responded to the FRA protocol after task engagement, see **Supplementary Table 7**). We found that the latency to peak was slower in adolescent AUDp. The FWHM and minimal latency were slower in the adolescents in both AUDp and AUDv. In the adolescent AUDv, additional parameters (like spontaneous- & evoked-FR, trial-responsiveness, life time sparseness), were different than those in adults. There were no significant differences between AUDp and AUDv within age-groups (**Supplementary Table 8**).

Next, we tested the learning-related effects among the different groups. The peak of frequency tuning was heterogeneous across the frequency range with no clear trends of learning or age (**Fig. 7d**). To quantify these responses, we calculated tuning width (**Fig. 7e**), population sparseness (**Fig. 7f**), the distance between the best frequency of the neurons and the Go frequency of the easy task (**Fig. 7g**), and single neuron discriminability (**Fig. 7h**). As expected from previous work, learning significantly changed (3 out of 4) neuronal tuning properties in adults (**Fig. 7e-h**, ‘Adults’). Surprisingly, however, learning had no significant effects on any of these single neuron parameters in the adolescent group (**Fig. 7e-h**). To test this at specific brain regions we analyzed AUDp and AUDv separately, which revealed specificity in learning-induced changes. Learning significantly affected 3 out of the 4 parameters in AUDp of adolescents (**Fig. S7-1a-d**). However, neurons from AUDv of adolescents showed no learning-induced changes (**Fig. S7-1e-h**). Notably, adult learning-induced plasticity was stronger than that in adolescence in all parameters besides the ‘BF to Go’ (**Fig. S7-1**). Thus, despite the notion that adolescents’ brains may be more plastic

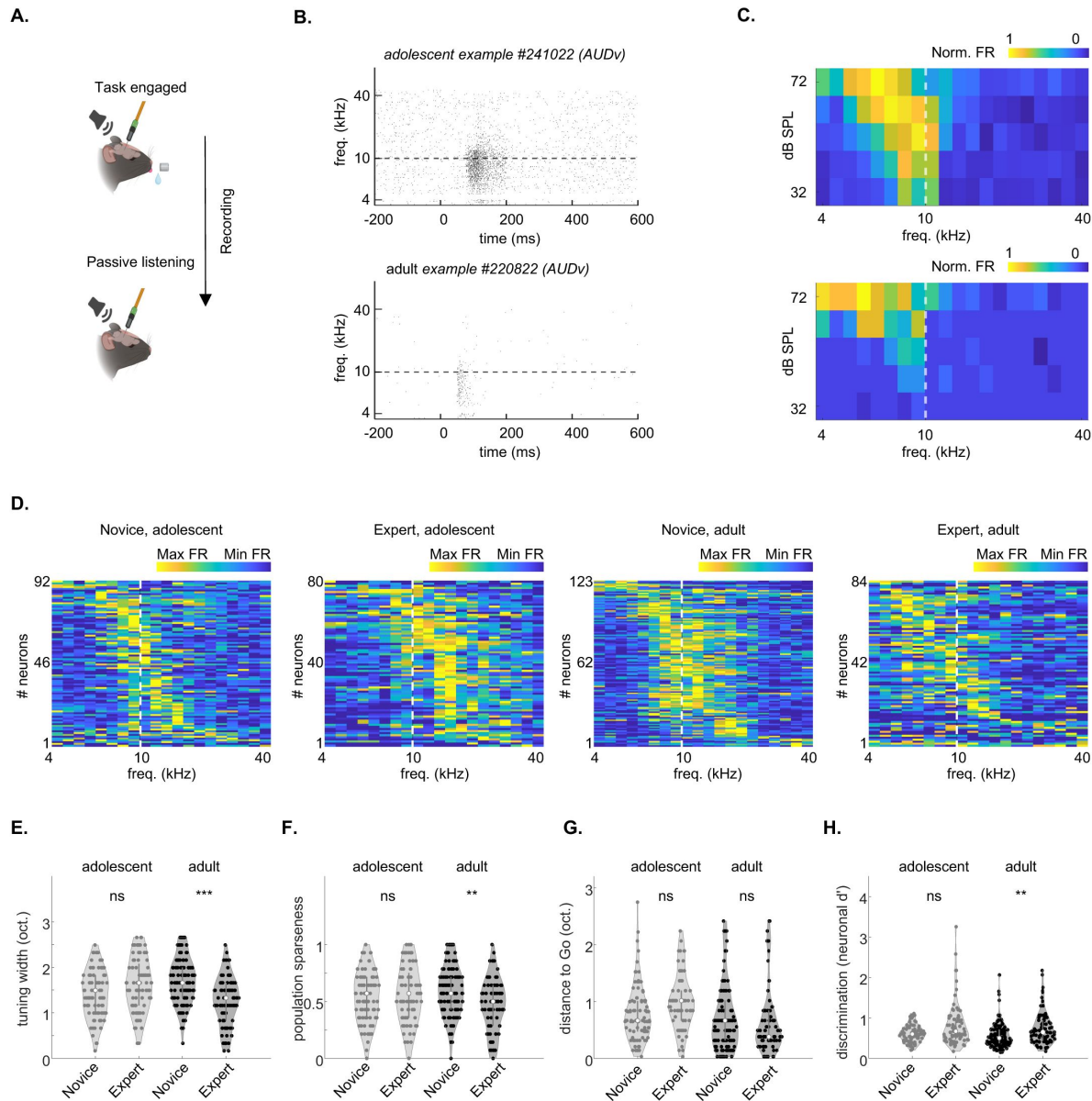


Figure 7

Adult mice show greater plasticity after learning.

A. Schematic showing that for the passive listening protocol, we continued our recording following the session of the engaged task (i.e. in satiated mice) by removing the waterspout. Created with *BioRender.com*. **B.** Example raster plot of a neuron from an adolescent mouse (top) and an adult mouse (bottom). **C.** FRA's of the neurons shown in 'B'. **D.** Distribution of best frequencies in our dataset. Values are normalized firing rates calculated at 62 dB SPL. Matrices are sorted by BF for clarity. Dotted line marks the decision boundary. **E.** Tuning bandwidth at 62 dB SPL of neurons in adolescents and adults. Side by side comparisons of novice versus experts. (adolescents $p = 0.0882$, adults $p = 0.0001$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **F.** Same as E. for Population sparseness (adolescents $p = 0.9549$, adults $p = 0.0013$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **G.** Same as E. for the distance (in octaves) between the best-frequency of each neuron to the easy Go-stimulus (adolescents $p = 0.0816$, adults $p = 0.6391$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **H.** Same as E. for the average neuronal d' of frequencies in the learned frequency spectrum (adolescents $p = 0.1627$, adults $p = 0.0026$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons).

than adults, we found generally stronger learning-induced plasticity in basic tuning properties in adults. These results underscore that learning-induced plasticity manifests differently in adolescents and adults.

Discussion

The term “stormy adolescence” aptly captures what we know from psychology, neuroscience, and biology—that adolescence is a period of turbulent development, which shapes the individual’s future behavior and health[33, 34]. While numerous behavioral phenomena of adolescence have been studied, particularly in humans, the neural underpinnings of these remain largely unknown[35]. In this study, we focused on (auditory) cortical neurons of the adolescent brain of mice while they were engaged in an auditory discrimination task. We found numerous differences in behavior, single neuron activity, and population encoding between adolescents and adults.

Adolescence behavior – perceptual and cognitive differences

We utilized the Educage platform to compare between adolescents and adults[18]. The Educage provides an optimal environment for training young mice; its automated nature minimizes human handling, which increases efficiency and reduces to a minimum any interference during the training process. Using the Educage, we found that procedural learning was intact in pre-adolescent mice (P23-P30), and behavioral differences were found only after the mice learned the task during the hard discrimination phase (P30-37). Compared to adults, adolescent mice scored lower on the hard versions of the task (Fig. 1i-j). Interestingly, our data suggests that the lower performance in the hard task in adolescents was not due to a deficiency in perceptual sensitivity (Fig. 2c), but rather to a cognitive deficit – namely, lick bias (Fig. 2d). This result is perhaps not surprising since higher lick bias is a cognitive phenomenon that has been well documented in adolescence as a maturation period of inhibitory control[36]. Indeed, additional cognitive control mechanisms are still developing during this age [37–39].

Another manifestation of weaker cognitive control mechanisms is the noisy behavioral performance of adolescents (Fig. 2h-j). While this noise may reflect higher plasticity and potentially serve cognitive flexibility[6, 40], it poses a disadvantage for the precision and consistency required in our task. This noise may also arise from changes in pubertal hormones, which we did not manipulate[41]. Notably, few studies have explicitly characterized auditory discrimination of adolescents in rodents; the majority of these have been carried out in gerbils using an amplitude modulation (AM) detection task [4, 26, 42]. Despite the different task, many of our age-related findings are aligned with the findings in gerbils, including lower performance in adolescents, and greater variance in adolescent behavior[4]. However, some findings differ. For example, the lick bias which was different in our data, was similar between young and adult gerbils in an AM modulation detection task. This difference may arise from differences in task design, as well as the severity of the punishment used in the task (here, we used a milder punishment). Here, our data support a view that cognitive factors, and to a lesser extent perception of pure tones *per se*, are the primary contributors to the poorer performance observed in adolescent mice when performing a pure-tone sound discrimination task.

Representation of stimulus and choice in ACx still undergoes maturation in adolescence

Motivated by the abovementioned behavioral differences, we tested how single neurons in ACx represented sounds (by comparing Hits vs FA) and choices (by comparing FA to CR). While cortical representations of sounds in ACx could be the basis of perceptual constraints, distinct processing of choices would reflect that more cognitive mechanisms are involved. We found that cortical representations of sounds and choices in mice engaged in the task were still immature as

compared to adults, e.g., responses in adolescents were more delayed and less informative (**Figs. 4–6**). These differences can be a result of age-related differences, and/or the result of a combined ‘age x learning’ effect.

We found only modest age-related differences in spiking activity of layer 5/6 neurons between adolescents and adults (**Fig. 7**). This limited difference is perhaps not surprising because the CP for tonal organization in ACx is already complete by the age at which we tested the animals (i.e., on P37-P42 [9, 13]). In contrast, more pronounced differences emerged from the interaction of ‘age x learning’, specifically in mice that learned the task and were actively engaged in the behavior (**Fig. 4f**; **Fig. 6d–g**). This interaction is evident in the lower single-neuron discriminability (AUC) during the hard task in expert adolescent mice (**Fig. 6d vs. 6f**).

While this may seem counterintuitive, it is compatible with the idea that learning can drive complex reorganization in sensory representations. For example, we recently showed that learning leads to distinct and sometimes opposing effects on single-neuron responses with respect to whether the task is perceptually easy or difficult [46]. It is also important to note that our recordings were restricted to infragranular layers (5/6) and did not sample supragranular (layers 2/3) neurons, which may exhibit distinct learning related plasticity [21]. As such, the developmental differences we report may reflect layer-specific properties of the deep output layers.

We speculate that immature feedback to ACx may underlie these differences for several reasons. First, feedback from higher cortical regions to more primary cortices has been proposed as a neural substrate for difficult perceptual discriminations [47, 48]. Second, feedback activity (often referred to as ‘top down’) is thought to be the key neural pathway involved in cognitive control mechanisms [49]. Feedback is typically expressed in the late (>100ms) phase of single neurons activity, which was significantly different between adults and adolescents across our measured parameters (onset-latency, duration of discriminability and maximal discriminability; **Fig. 4g**, **Table 2**, **Table 3**). The precise source of feedback that contributes to perceptual or cognitive differences remains to be determined, but likely candidates include regions that project to ACx [50, 51], such as medial prefrontal cortex [52, 53], orbitofrontal cortex [54], or anterior cingulate [55]. Testing the causal role of these regions during adolescence will be an important direction for future research.

Cortical coding is performed by populations of neurons, which can efficiently encode stimuli on a trial-by-trial basis [56–60]. Population coding, as measured here by decoding accuracy, largely recapitulated the results of the single neuron data. Specifically, cortical populations in adults encoded the different tones better and faster; particularly in the hard task (**Fig. 5a–c**). Interestingly, by testing the quality (i.e., separation) of the decoding performance, we revealed a correlation between decoding in the easy and hard tasks in adult (**Fig. 5d**). While this result reflects the collapsed information of stimulus and choice, it suggests that the cortical network uses shared representation for decoding the easy and hard tasks, possibly by the same neurons. This is consistent with our recent work showing that indeed the same neurons in ACx respond to the same stimulus differently, depending on if the mouse is engaged in discriminating an easy sound pair or a hard sound pair [46]. The shallower slopes in adolescence suggest a less efficient decoding scheme that may arise from different populations of neurons encoding task-related information.

Our optogenetic silencing of ACx in expert adult mice resulted in a marked suppression of licking behavior (**Fig. S3-2**), consistent with the idea that ACx contributes to task performance. However, the strength of the observed effect raises the possibility that the manipulation may have impacted broader aspects of behavior—such as arousal, motivation, or motor readiness—rather than selectively disrupting auditory discrimination. Recent work [61] suggests that ACx plays a critical role during learning by supporting higher-order computations but may become less

essential for task execution once mice reach expert performance. These findings, together with previous reports [19, 20], indicate that the necessity of ACx for auditory-guided behavior may depend on task structure, training stage, and the exact nature of the manipulation. The robust decrease in lick rate and performance observed in our experiment may reflect the intensity, specificity, and timing of the silencing. The use of GtACR2, a highly potent silencer, may have further amplified the effect. We cannot rule out the possibility that silencing the output from ACx also disrupted the flow of information to downstream areas, thereby exerting a stronger influence on perception and behavior. Although the injection was targeted to the ACx, off-target effects are also likely to have contributed. Future experiments using graded or temporally specific perturbations will be useful to more precisely define the contribution of ACx across learning and development.

Learning in the adolescent brain

Cortical representations of pure tones have a critical period (CP) between P12-P15 [13, 16, 17]. Indeed, we chose to study animals at P37-P42 on a pure tone discrimination task precisely to avoid the hypersensitivity associated with this CP. But this may be unavoidable, as recent work has shown that ACx also has a later CP for more complex sounds features [16, 17]. Those late CP's overlap with our training schedule and the underlying biological manifestation of this CP is expected to affect learning related plasticity in general. The underlying mechanism of the late CP to complex sounds has been associated with dynamic changes of inhibitory circuits [62, 63]. Thus, immature inhibitory circuits remain sensitive in adolescents, and could affect how pure tones are represented, despite the closure of the classic tonal CP. Notably, sound discrimination learning changes tonal representations even in adult mice, well after the closure of CP's for both simple and complex sounds [18, 24, 46]. We hypothesize, therefore, that the differences in performance on more difficult discriminations between adolescents and adults are shaped primarily by age-specific learning-related plasticity, rather than by differences in pure tone representation per se.

Taken together, our findings show that cortical responses in adolescents differ from those in adults, despite both groups learning the same pure tone discrimination task under identical training protocols. Although the CPs for pure tones are considered closed by this age, learning to discriminate between tones still varies across developmental stages. It remains unclear how later CPs, such as those shaping responses to amplitude modulation or frequency-modulated sweeps through inhibitory circuits, contribute to these differences. Likewise, the role of ongoing maturation in cognitive control, including the development of feedback pathways, warrants further investigation. Nevertheless, it is evident that developmental processes in the adolescent brain continue to influence how even simple sounds are encoded through learning. The adolescent-specific plasticity observed here directly reflects on how perceptual and cognitive features are integrated within the auditory cortex.

Methods

Animals

A total of 47 (37 female and 10 male) C57BL/6 mice were used in this study. Adolescent mice were weaned at postnatal (P) day 20. Adult mice were trained starting from P60. All mice were housed on a 12-hour light/12-hour dark cycle with food and water *ad libitum*, unless used for behavioral training (see below). All experiments were approved by the Institutional Animal Care and Use Committee (IACUC) at the Hebrew University of Jerusalem, Israel (Permit Number: NS-21-16448-4, NS-21-16694-4, NS-22-16966-4).

Surgical procedures

The details for surgical procedures were identical to those used in previous studies from our laboratory [18, 22, 50, 64–67]. The relevant procedures are briefly described below.

RFID Implantation

24h before the start of behavioral training in the automated home-cage, mice were implanted with a radio frequency identification (RFID) chip (Trovan, EXTM). Animals were anesthetized with 2% Isoflurane with pure O₂ as carrier. RFID chips were implanted under the scruff. After implantation, mice were injected with 0.04mg/g of 10% Meloxicam solution to prevent infection and alleviate any possible pain from the implantation.

Head-bar Implantation & Craniotomy

48h before the start behavioral training on the head-fixed recording setup, mice were anesthetized with 2% Isoflurane, the body temperature maintained at 37 °C and the eyes were covered with 5% Chloramphenicol ointment to prevent from drying. Before removal of the scalp, we applied 4% Lidocaine above the skin and skull area. Afterwards, we glued a custom-made titanium bar at the back of the mouse skull using dental cement (Meta-bond). To enable acute recordings in ACx, we performed a craniotomy on the left-hemisphere at 2.5 mm posterior and 4.0 lateral to the bregma. The craniotomy was protected by a pool of dental cement and covered with a silicon elastomer (WPI; Kwik-Cast catalog #KWIK CAST). After the craniotomy, mice were injected with 0.04mg/g of 10% Meloxicam solution and 0.2 ml Saline. During the 48 hours of recovery, mice received two further doses of 0.04mg/g of 10% Meloxicam solution. 24h before recordings, mice were again anesthetized to remove regrowth of the dura-mater during training. The area around the craniotomy was cleaned with Hydrogen-peroxide (3%), the silicon elastomer was replaced, and another dose of 10% Meloxicam was injected.

Virus Injection & Cannula Implantation

The surgical procedure was similar to the *head-bar implantation & craniotomy*. The differences in the procedure are outlined below. Bilateral craniotomies were performed at 2.5 mm posterior and 4.3 mm lateral to the bregma. We injected 200nl AAV5-CAMKII-GtACR2-FRED_kv_2.1, or 200nl AAV9-CAMKII-dTomato. Both viruses were produced at the ELSC virus-core facility. We injected the virus at a depth of 1.1 mm at 0°. Afterwards, we inserted a 0.2 mm diameter optical fiber with an attached cannula (CFML1202; Thorlabs) at a depth of 0.9 mm at 0°, bilaterally. We chronically fixed the fiber position on the skull using dental cement (Meta bond).

Auditory Stimuli

Sound stimuli were delivered from a calibrated free-field speaker (ES1 SN:4568) using a multifunction processor (ED1; Tucker-Davis Technologies). The speakers were calibrated with a free-field microphone (Type 4939, Bruel and Kjaer). The stimuli were comprised of pure tones, sampled at 500 kHz, between 7.07 to 14 kHz for behavioral training and 4-40 kHz for passive listening. All stimuli were played at sound pressure levels of 72 dB SPL for behavioral training, and 72-, 62-, 52-, 42-, and 32-dB SPLs for passive listening. All tones were played in randomized order.

Behavior

Automated home-cage training (the ‘Educage’)

For a detailed description of the EduCage behavioral platform, see [18]. After RFID-implantation, groups of 5 mice were placed in their home-cage that was connected to a behavioral chamber called the Educage (Fig. 1a). Mice were provided with food *ad libitum*. Water access was restricted to the Educage chamber. Mice could access the Educage freely to retrieve water mixed with 5% sucrose. The behavioral training was controlled by a custom LabView 2019 (National Instruments) program running on BNC-2110 (National Instruments) and FPGA (MyRIO – National Instruments) to register licks, RFIDs numbers, and deliver stimuli and reinforcements. Every time a mouse entered the water port its RFID was automatically recognized, a trial initiated, and an auditory stimulus played from a speaker vertically positioned above the behavioral chamber. We trained mice gradually for three weeks on a Go/No-Go paradigm of pure tones, spaced around a category boundary of 10 kHz. This paradigm enabled us to efficiently compare between adolescent and adult behavior, since learning was restricted from post-weaning (P20) to adolescence (P37) (correspondingly adults were trained from P60-P77). Mice were trained in four separate training phases (Fig. 1B). For the first 24h (P20-P21 in the adolescent group of mice and P60-P61 in adult mice) mice were habituated to the behavioral chamber and could approach the port freely to retrieve water after licking the waterspout (0.0015 ml per trial). During the following 48h (P21-P23 in the adolescent group of mice and P61-P63 in adult mice) mice underwent a tone-association phase. All tones played were 300ms long (sampled at 500 kHz). Tone association was restricted to the lowest Go tone (7.07 kHz at 100% probability). Mice received a water-reward in response to licking during the first 2 seconds after tone-offset. This was followed by the first tone-discrimination stage, which contained a No-Go tone (14.14 kHz at 45% probability) at 1 octave away from the Go tone. We analyzed four response outcomes—Hits, Misses, Correct Rejections, False Alarms—in the 2 seconds response window after the tone-offset. ‘Hits’ were counted after 5 successful licks to the Go-tone and followed by a water-reward. ‘Misses’ were counted when licks to the Go-tone trials did not pass the lick threshold. Correspondingly, ‘Correct Reject’ (CR) trials, were all No-Go trials below the lick threshold. CR trials were not reinforced. ‘False Alarms’ (FA) were counted when 5 or more licks were registered after the No-Go. FA trials were followed by a white-noise punishment (5-20 kHz at 72 dB SPLs, 2 seconds) and an inter trial delay of 5 seconds. In addition, mice were exposed to 5 probe trials (tones whose trial outcome that were neither rewarded nor punished at 8.49 kHz, 9.567 kHz, 10 kHz, 10.44 kHz, and 11.89 kHz; played at 2% per tone). Mice trained on the 1 octave tone discrimination for 1 week (P23-P30 in adolescent mice and P63-P70 in adult mice). Next, we added a second pair of tone - a Go/No-Go tone pair spaced 0.25 octaves apart (9.17 kHz and 10.95 kHz) for another week of training. All tones were played in randomized order and trial difficulties of the easy and the hard task were intermingled. The experiment was terminated at age P37 for adolescent and P77 for adult mice.

Head-fixed training

We used a head-fixed paradigm that enabled us to efficiently train adolescent and adult mice on the same task and record from engaged mice after learning. We trained mice on the same Go/No-Go stimulus setup of the Educage described above with few differences which are outlined below.

The head-fixed behavioral training was controlled by a custom MATLAB 2023B (Mathworks) program running on BNC-2110 (National Instruments) to register licks and deliver stimuli and reinforcements. Mice were water-restricted for 24h prior to training, and their weight monitored daily. While training on the setup, mice received water supplemented with 5% sucrose, depending on their weight (0.125 ml per g of body weight). Additional water, in the form of Hydrogel (ClearH2O) was given after daily training if mice did not consume the required amount during training, or their body weight dropped below 85% of the pre-restriction weight. Mice were gradually habituated to the setup by head-fixation and received free water after licking, during the first training day (P23 in adolescent mice and P63 in adult mice). Licking was measured with an optical lick-meter (Sanworks). During the second training day mice were associated with a tone.

All tones of the task were 100ms long (sampled at 500 kHz). Mice received a water reward in response to licking in the first 2 seconds after tone offset. Reward-reinforcement was delayed to 0.5 seconds after the tone offset. To break the regularity of the trial sequences, we applied a dynamic inter-trial interval of 6-8 seconds. New trials were initiated if mice did not lick the spout at least 2 seconds prior to the upcoming trial. Mice were familiarized with the tone for at least two training days (P24-26 in adolescent mice and P64-66 in adult mice). Then, and when the lick rate was above 80%, mice proceeded to the first tone discrimination stage. We gradually increased the probability of No-Go trials, depending on the lick rate of the mouse. We added the 0.25 octave Go/No-Go tone pair, after a maximal of seven training days on the 1 octave tone pair (P26-33 in adolescent mice and P66-73 in adult mice), or if the behavioral performance exceeded a threshold of $d' = 1$, as calculated using the signal detection metric. Similar to the Educage, all tones were played in randomized order and trial difficulties of the easy and the hard task were intermingled. Mice were trained until P37 in the adolescent group and P77 in the adult group.

Optogenetic Manipulation

Adult mice (P60) were injected with AAV5-CAMKII-GtACR2-FRED_kv_2.1 ($n = 3$) and implanted with cannulas prior to training as described above. Mice underwent optogenetic stimulation after training on the head-fixed task by attaching a ferrule patch cable (M79L01; Thorlabs) to the implanted cannula using a mating sleeve (ADAF1; Thorlabs). The ferrule cable end (SMA connector) was connected to an LED driver setup (LEDD1B; Thorlabs) that enabled us to send precise outputs at 476 nm (blue light; 5mW) bilaterally. To test the effect of ACx inhibition on task performance, we timed the LED output to 50ms before tone onset until 50ms after tone-offset. The LED consisted of pulses given at 10 Hz for 200ms. Behavioral sessions of optogenetic manipulation were identical to the Go/No-Go paradigm of previous training sessions. We inhibited ACx in 50% of the trials in randomized order. Optogenetic manipulations were repeated throughout multiple sessions of expert performance. As a control, we repeated the same experiment in mice injected with AAV9-CAMKII-dTomato ($n = 3$). After completion of the experiment, animals were deeply anesthetized with Ketamine and medetomidine (0.008 g/kg, and 0.65 mg/kg, respectively) and then perfused with 4% Paraformaldehyde. After the perfusion, the brain was extracted and preserved. We then sectioned the ACx using a cryostat (Leica) into 0.05 mm thick coronal slices. Afterwards brain slices were washed one time under 1% Phosphate-buffered saline (PBS), and a second time with 1% PBS plus 0.4% Triton. Finally, slides were mounted and stained with DAPI (4%) and imaged using a wide-field fluorescent microscope (Olympus Life Science, Olympus IX83), to reconstruct the probe position offline.

Extracellular Recordings

Recording setup

Before the onset of the recording session, we removed the silicon elastomer and placed an external reference electrode (Ag/AgCl wire) on the skull of the right hemisphere. All recordings were performed using Neuropixels 1.0 (Npx; IMEC, phase 3A), together with a base-station connected to a chassis (IMEC; NI PXIe -1071, National Instruments). Probes were mounted to a custom-made steel rod and connected to the ground. Before penetration, probes were covered with a fluorescent dye (Dil (Invitrogen catalog #V22885-red, or Dio (Invitrogen catalog #V22886 – yellow), to reconstruct penetration sites after the recording.

Then, Npx probes were inserted into the left ACx at 30° to a depth of 3.85 mm (this approach limited our recording to deep layers of the ACx). The skull surface was submerged in saline and the probe allowed to settle for 10 min. We sampled recordings at 20 kHz, with action potential band filtered to contain 0.3-to 10kHz frequencies. Action potential band gain was set to 500. Out of

the 960 available sites on the 1 cm shank of the Npx probe [68], we acquired the 384 lowest recording shanks in a staggered configuration. We used common-average referencing to process the voltage-traces.

Recording Schedule

We performed recordings of both novice and expert adolescent and adult mice. Novice recordings were performed after tone-association (see *head-fixed training*; adolescent $n=3$; adult $n=3$). Expert recordings were performed after learning (adolescent $n=5$; adult $n=6$). Both novice and expert recording include the easy (1 octave) and hard (0.25 octave) tone pair. Both recordings were performed after P37 for adolescent mice and after P77 for adult mice. Tone probabilities and protocol parameters were identical to the *head-fixed training*. Each mouse was recorded multiple times (novice recordings adolescents $n=6$; novice recording adults $n=6$; expert recordings in adolescents $n=13$; expert recordings in adults $n=14$). In some mice, after the recording in the engaged configuration, we recorded neural activity under passive listening conditions to a pure tone protocol (novice recordings adolescents $n=6$; novice recording adults $n=6$; expert recordings in adolescents $n=4$; expert recordings in adults $n=4$). The pure tone protocol was comprised of 20 different frequencies, logarithmically spaced between 4 and 40 kHz, presented at 5 sound pressure levels (see *Auditory stimuli*). Each frequency and attenuation combination were presented 16 times. The tone interval was set to 1 second. After completion of the recording schedule, animals were deeply anesthetized with Ketamine and medetomidine (0.008 g/kg, and 0.65 mg/kg, respectively) and perfused with 4% Paraformaldehyde. After the perfusion, we performed histological analysis and sectioned coronal brain slices as described above (see *Optogenetic Manipulation*) to reconstruct the exact probe position.

Preprocessing & Spike-sorting

All recordings were sorted using Kilosort 2.5/3 open-source software ([69]; <https://github.com/MouseLand/Kilosort>). After automatic sorting, we performed manual curation of the acquired units using ‘Phy-2’ open-source GUI. (UCL; <https://github.com/cortex-lab/phy>). During manual curation we distinguished between single units (SUs) and noise (pre-labeled multi-units in Kilosort2.5/3 were automatically labelled as noise). The following parameters were set to determine SUs: physiologically plausible waveform shape (1–2 ms, biphasic/triphasic), high spike-amplitude ($> 50 \mu\text{V}$, $\text{SNR} > 3$), physiologically plausible refractory period ($\text{ISI} > 4\text{ms}$), sufficient inter-spike-interval (ISI peak 5–10 ms), sufficient firing-rate across the recording (> 0.1 spikes/sec), and principal-component cluster density (high density and low overlap). In addition, SUs were compared to neighboring units, based on waveform, firing rate, drift-pattern and cross-correlograms to determine merging of two SUs into a single cluster (Pearson Correlation Coefficient > 0.8). Units that passed the above-mentioned criteria were considered single units (SUs).

Data Analysis

The analysis was performed with MATLAB R2023b (MathWorks) and Python3.13 (PyCharm 2024.1.3). Violin plots are presented with the median (white dot) together with the interquartile range. Boxplots are presented with the median (horizontal line), mean (grey dot) and quartiles (box). Temporal plots are presented as the mean (line) and the standard error of the mean (patched or errorbar).

Behavioral analysis

We categorized behavioral responses of the task into hit, miss, false alarm, and correct reject responses. To compare behavioral performance, we calculated the d' value based on the signal detection metric ($d' = \text{the inverse normal distribution of the z-scored hit rate} - \text{the z-scored (false alarm rate)}$). Z-scored values of 1 were rounded to 0.99, to avoid d' approaching infinity. Unless,

defined differently d' was defined for the last 100 trials. Next, we calculated psychometric curves based on the lick rate (above the lick threshold) to learned and probe trials (Fig. 2b, c, e). Psychometric curves were normalized and fitted to a sigmoidal function, defined as follows:

$$S(t) = a \frac{1}{1 + e^{-\frac{(t-b)}{c}}} \quad (\text{Equation 1})$$

Where a denotes the lick rate, b = the time of the inflection point, and c = the steepness of the curve. Based on the normalized psychometric curve, we extracted two inflection points (random lick rate at 0.5 and category boundary at 10 kHz). To compare the heterogeneity of adolescent and adult behavior, we calculated the criterion bias (the inverse normal distribution of the $(-0.5 * (z\text{-scored (hit rate)} + z\text{-scored (false alarm rate))))$ and analyzed the coefficient of variation (CV) of behavior across the learning. We applied similar behavioral measurements to the Educage training and head-fixed training, as well as recordings and optogenetic manipulation during task performance. During expert-performance in the head-fixed task, we analyzed the running d' as the average performance of bins of 25 trials (Fig. 3). In addition, we also calculated the average lick trace during the reinforcement delay. Here, we also extracted the lick latency per trial per mouse, defined as the lick latency from tone-onset (Fig. 3).

Neuronal analysis

Probe trajectories and location of SUs were reconstructed using Allen CCF tools (<https://github.com/cortex-lab/allenCCF>; [70]). For further analysis we only considered SUs that were recorded from infragranular layers 5 and layer 6. SUs from the dorsal-auditory cortex (AUDd), primary auditory cortex (AUDp), ventral auditory cortex (AUDv) and temporal association cortex (TEa) were pooled together as auditory cortex (ACx), and in some analyses presented separately [65]. Neurons were considered excitatory auditory responsive if the average spontaneous firing rate (FR) of the baseline activity preceding the tone (200ms to 50ms before stimulus onset) was significantly lower than the average tone-related activity (from stimulus onset until 50ms after stimulus offset) across all trials [22]. The difference was tested with a right-sided Wilcoxon sign rank test ($p < 0.05$). Suppressed auditory responsive units and non-auditory responsive units were excluded from the analysis. Peri-stimulus time histograms (PSTHs) were smoothed with a gaussian smoothing filter of 5ms. To test the difference between adolescent and adult auditory firing properties in expert mice, we calculated the difference in average spontaneous FR (−200ms preceding the tone onset until tone-onset, average evoked FR (tone onset until 50ms after tone offset) and fraction of responsive trials. We also compared the coefficient of variance of the evoked FR (mean FR divided by standard deviation of the FR across trials). Temporal differences in firing properties were compared by calculating the latency to peak (latency of the highest FR), minimal latency (first spike after tone-onset), full-width-half-maximum (FWHM; time from peak FR to baseline FR), and the lifetime sparseness, which was calculated as follows:

$$S = \left(1 - \left(\frac{\sum r_i^2}{n} \right) \right) / \left(1 - \frac{1}{n} \right) \quad (\text{Equation 2})$$

r_i = corresponds to the FR of each learned frequency and n equals the number of learned tones. Values of S near 0 correspond to a dense FR and values near 1 to a sparse code [71].

Single neuron analysis during task performance

We investigated the FR of excitatory auditory responses per trial from 200ms preceding the stimulus up to 600ms after the tone onset. To test how well single neurons discriminate between trial outcomes, as well as task difficulty, we calculated task-related activity per trial using receiver operating characteristics (ROC) and calculated the area under the curve (AUC) across a running window of 25ms in bins of 50ms (Fig. 4, Fig. 6). Stimulus-related activity was defined as the

hit compared to the false alarm trials, separately for the 1 octave and 0.25 octave tone pair. Stimulus-related activity of miss compared to correct trials were not analyzed due to the lack of miss trials in expert performance. In addition, we also tested the average (mean of 1 octave and 0.25 octave tone pair) choice discrimination, defined as the difference between false alarm and correct reject trials (**Fig. 4**). We sampled 20 trials per trial outcome and repeated the ROC encoding 100 times. Afterwards, we adjusted the average AUC values of all iterations of each neuron to its baseline AUC before the tone onset ((AUC – baseline AUC) + 0.5). AUC values close to 0.5 indicate a low neuronal discriminability, and values up to 1 indicate high neuronal discriminability [72]. To test significant discrimination, we repeated the ROC encoding with shuffled trial identities (20 randomly shuffled trials per trial outcomes across 100 iterations). Significant discriminability of a neuron was defined as the AUC time bins that exceed the mean ± 3 std of the shuffled distribution [64]. Adolescent and adult neuronal discriminability were compared using the onset-latency of discrimination of a neuron (first timepoint > mean ± 3 std of the shuffled distribution), the maximal AUC of the running window and the duration of discriminability (time duration > mean ± 3 std of the shuffled distribution). To compare novice and expert performance we repeated the AUC encoding and compared the general discrimination of 1 octave and 0.25 octave Go and No-Go discrimination of all trial outcomes (**Fig. 6**). To evaluate the correlation of behavioral performance and neuronal discriminability, we compared d' per recording to the average of the maximal AUC of all neurons per recording. The AUC and d' values were compared via pairwise Pearson correlation.

Auditory cortex population analysis during task performance

We analyzed the population activity in the ACx of adolescent and adult mice on a trial-by-trial basis. Population activity was segregated by hit and correct-reject trials for both easy and hard tasks in each age group. We excluded recordings that contained fewer than 20 auditory-responsive neurons, fewer than 20 hit or correct-reject trials in either task. This resulted in a total of 8 adolescent sessions and 10 adult sessions. To decode the stimulus from neural activity, we applied linear-discriminant analysis (LDA) to the first 200ms following stimulus onset. Decoding accuracy was calculated on held-out trials (10 trials per stimulus; **Fig. 5a**). To examine changes in decoding accuracy over time, we used a 50ms bin width and fitted separate LDA models at each time point, covering a period from –0.5s to 10s after stimulus onset (**Fig. 5c**). Next, we computed the separation per session, defined as the ratio of the between-class variance to the within-class variance. Given the means μ_1, μ_2 and variances Σ_1, Σ_2 of the two classes, the separation is given by:

$$S = \frac{\sigma_{between}^2}{\sigma_{within}^2} = \max_w \frac{(w(\mu_1 - \mu_2))^2}{w^T(\Sigma_1 + \Sigma_2)w}. \quad (\text{Equation 3})$$

We used robust linear regression without intercept (Huber loss) to fit the relationship between separations in easy and hard tasks (**Fig. 5d**). To quantify when the decoding accuracy is significantly different from the baseline, we used the bin width with a sample frequency of 100 Hz. Baseline decoding accuracy was calculated as the mean and variance of decoding accuracy during a 500ms window before stimulus onset, up to 50ms before onset. Decoding latency was defined as the time from stimulus onset to the first time point where decoding accuracy exceeded three standard deviations from the baseline mean (**Fig. 5b**). One adolescent session, in which behavioral d' is less than 1, has decoding accuracy that does not exceed three standard deviations from the baseline mean within the first 200 ms. Therefore, this session was excluded from the latency analysis. Finally, we assessed the single-trial variance across both age groups and tasks. Data were standardized per neuron before computing variance per stimulus (**Fig. 5e**). To visualize the results, we projected the population representations to the space spanned by the two LDA decoding dimensions for the easy and hard tasks. We averaged the mean and covariance matrix across sessions for each stimulus and age group, representing the distributions with corresponding ellipses (**Fig. 5f**).

Single neuron analysis during passive listening

We extracted the frequency-response areas (FRAs) based on the pure-tone responses (from tone onset up to 50ms after tone-offset) of auditory responsive neurons. The bandwidth of excited units was computed for significant excitatory responses of adjacent frequencies at 62 dB SPL and subtracted by the number of expected false-positive responses. Afterwards the sum of significant frequency responses per neuron was multiplied by the octave distance between every frequency to receive an octave-based bandwidth measure. We then calculated the population sparseness as the fraction of significant excited responses in adolescent and adult novice and expert mice. To test the neuronal discriminability to pure tone responses, we calculated the pairwise d' based on the FR per trial in all frequencies at 62 dB SPL [22, 66]. For two given frequencies p and q , d' values were calculated in the following way:

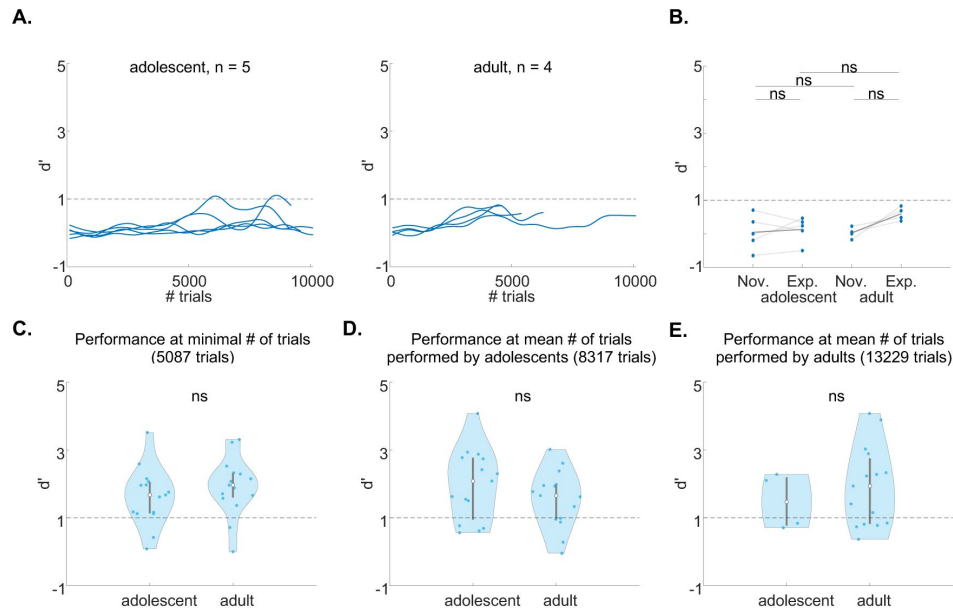
$$\frac{d(\mu_p, \mu_q)}{\mu(\sqrt{\sum_{i=1}^n (\hat{p}_i - \hat{q}_i)^2})} \quad (\text{Equation 4})$$

Where d represents the distance between the mean FR (μ) of frequency p and q , as vectors with n entries representing the mean signal (averaged over individual trials) in n -dimensional space for frequency p/q . The average distance (d) is then divided by the mean of the inner Euclidian distance ($\sqrt{\sum_{i=1}^n (\hat{p}_i - \hat{q}_i)^2}$) between each single trial (Fig. 7h).

Statistical Analysis

Statistical comparisons were performed in custom-written codes in MATLAB 2023b (MathWorks) and Python 3.12. We assessed if the data was normally distributed using a Kolmogorov-Smirnov test. Normally distributed data was tested using a paired (within group comparison) or independent (between group comparison) t-test and presented with mean $\pm$ STE. Non-normally distributed data was tested using a Wilcoxon sign rank (within group comparison), or rank-sum (between group comparison) test. Multiple-samples were tested with ANOVA (parametric data) and Kruskal-Wallis tests (non-parametric data). All tests corrected after multiple comparisons (Bonferroni-correction for two-samples and Tukey-Kramer for multiple samples). In addition, we applied linear-mixed-effect models (LME) to account for hierarchical data structures and variability within groups of co-housed mice, as well as repeated measurements of mice and recordings, as random effects. Fixed effects, and pairwise interaction effects were adjusted with post-hoc tests and multiple comparisons. For all statistical tests significant differences were defined as p-values below 0.05.

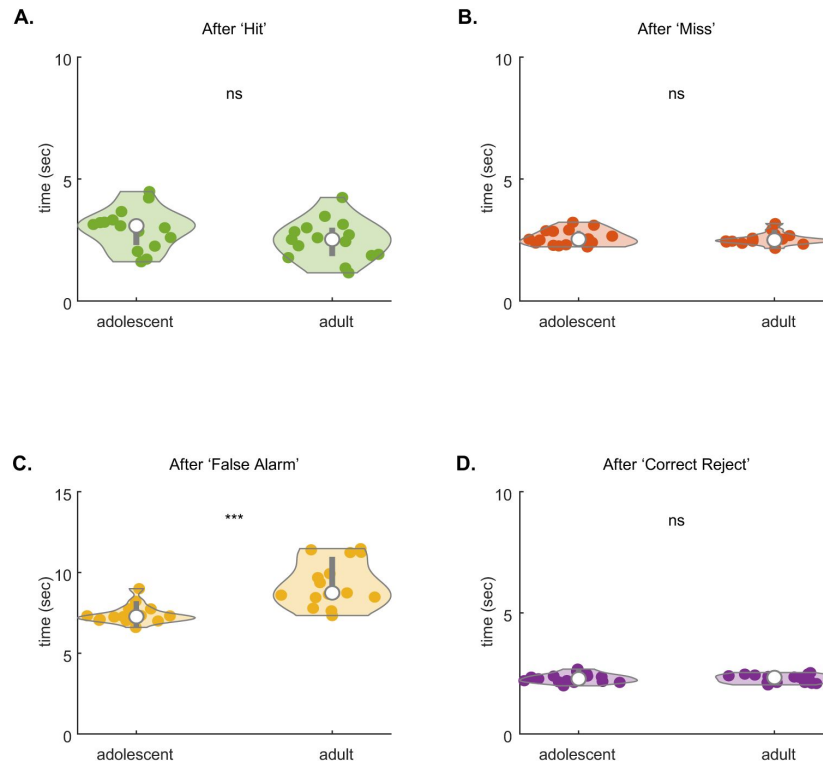
Supplementary figures and tables



Supplementary Figure 1-1

Behavioral criteria of auditory learning.

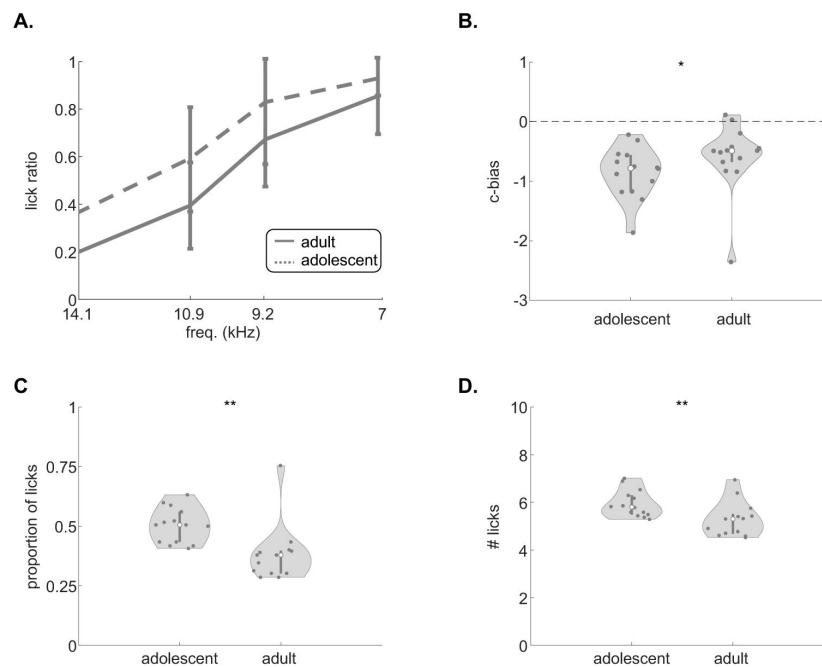
A. Learning curves per mouse throughout the experiment (n=5 adolescents, n=4 adults). **B.** Discriminability of novice (Nov.; first 100 trials) compared to expert (Exp.; last 100 trials) mice in the adolescent and adult groups (Novice vs Expert — Adolescents, $p = 0.0625$; Adults, $p = 0.0625$, Wilcoxon sign ranked test after Bonferroni correction; Adolescents vs Adults — Novice, $p = 0.4444$, Expert, $p = 0.5476$, Wilcoxon rank sum test after Bonferroni correction). **C.** Discriminability (d') of the easy task at the minimal number of trials (5087 trials) shared between all mice ($z = -1.0370$; $p = 0.2998$, two-sample Wilcoxon rank sum test). **D.** Discriminability (d') of the easy task at the mean number of trials of adolescent mice (8317 trials) shared between all mice ($n = 15$ per group; $z = 0.9125$; $p = 0.3615$, two-sample Wilcoxon rank sum test). **E.** Same as A. but for the mean number of trials (13229 trials) of adult mice (adolescent $n = 4$; adult $n = 15$; $z = 0.9125$; $p = 0.5304$, two-sample Wilcoxon rank sum test).



Supplementary Figure 2-1

Inter trial interval after different trial outcomes.

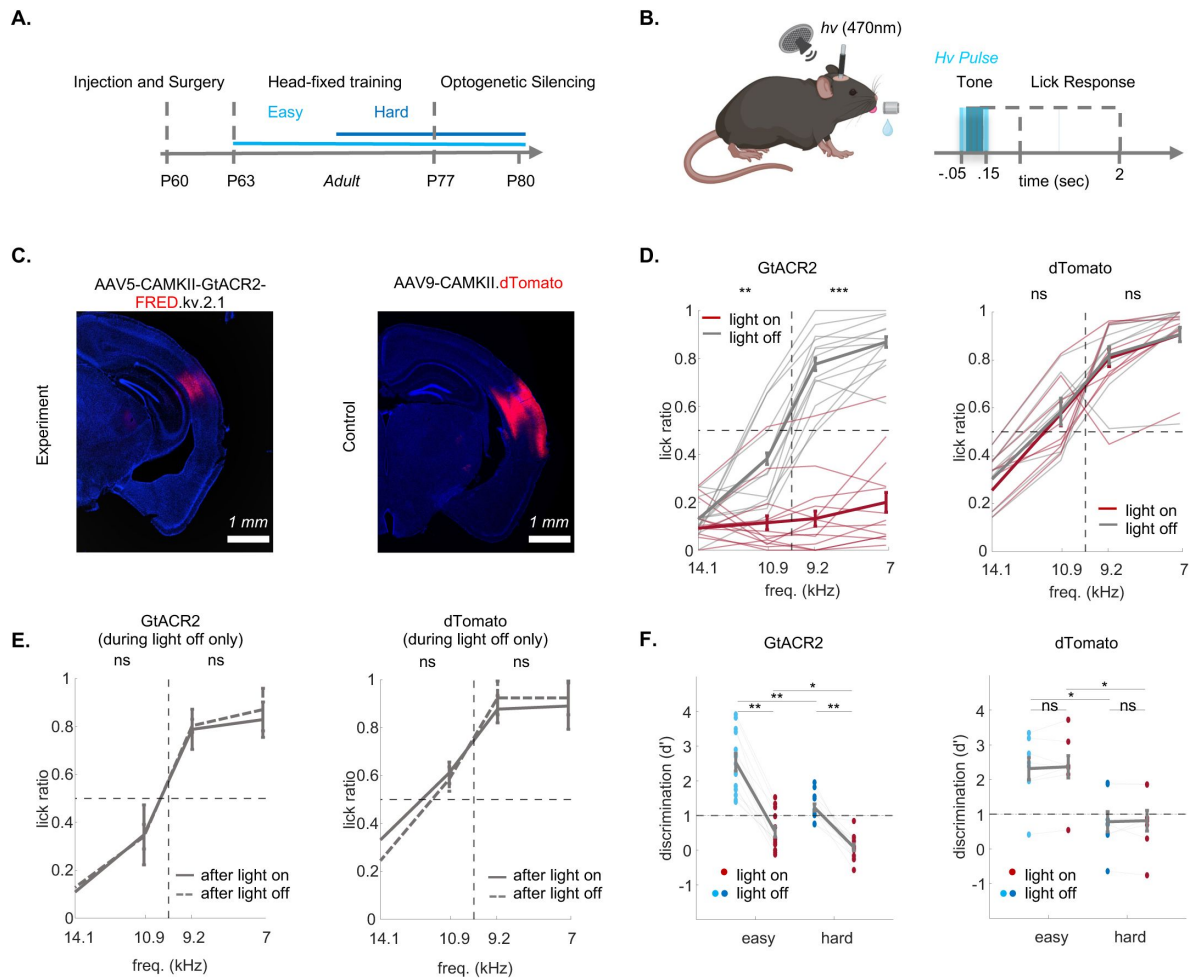
A. Average *inter trial interval* (ITI) per mouse to the next trial after a previous hit ($z = 1.6176$; $p = 0.1057$, two-sample Wilcoxon rank sum test). **B.** Same as A. after a previous miss hit ($z = 0.0830$; $p = 0.9339$, two-sample Wilcoxon rank sum test). **C.** Same as A. after a previous false alarm hit ($z = -3.9823$; $p = 6.8241 \times 10^{-5}$, two-sample Wilcoxon rank sum test). **D.** Same A. after a previous correct reject hit ($z = 0.3733$; $p = 0.7089$, two-sample Wilcoxon rank sum test).



Supplementary Figure 3-1

Lick bias and impulsivity in adolescent and adult mice during head-fixed recordings.

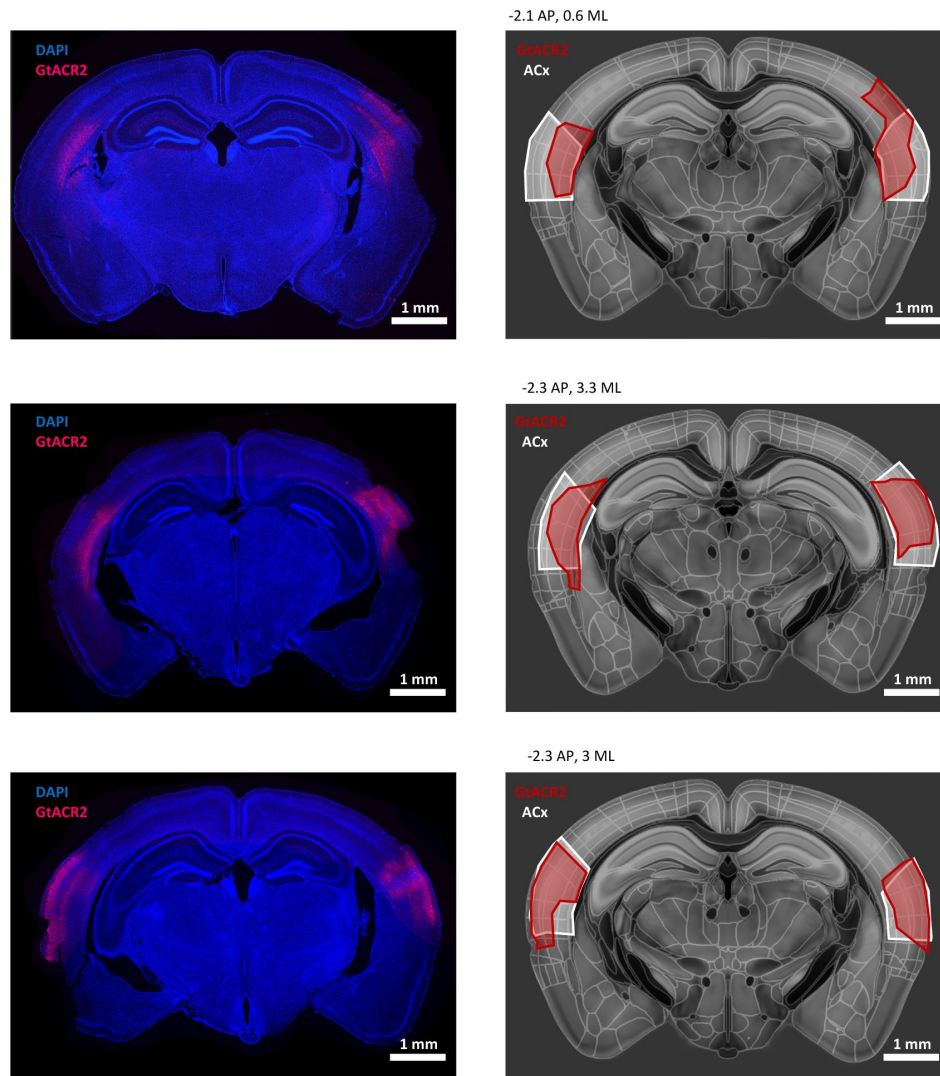
A. Average psychometric curve of adult recordings (solid line) and adolescent recordings (dashed line). **B.** Lick bias (i.e., criterion bias) per recording (c-bias: $z = -2.1366$, $p = 0.0326$, Wilcoxon rank sum test). **C.** Proportion of licks within ITIs after FAs ($z = -2.6447$, $p = 0.0082$, Wilcoxon rank sum test). **D.** Average number of licks during ITIs after FAs ($z = -2.7230$, $p = 0.0063$, Wilcoxon rank sum test).



Supplementary Figure 3-2

Auditory Cortex is necessary for task execution in adult mice.

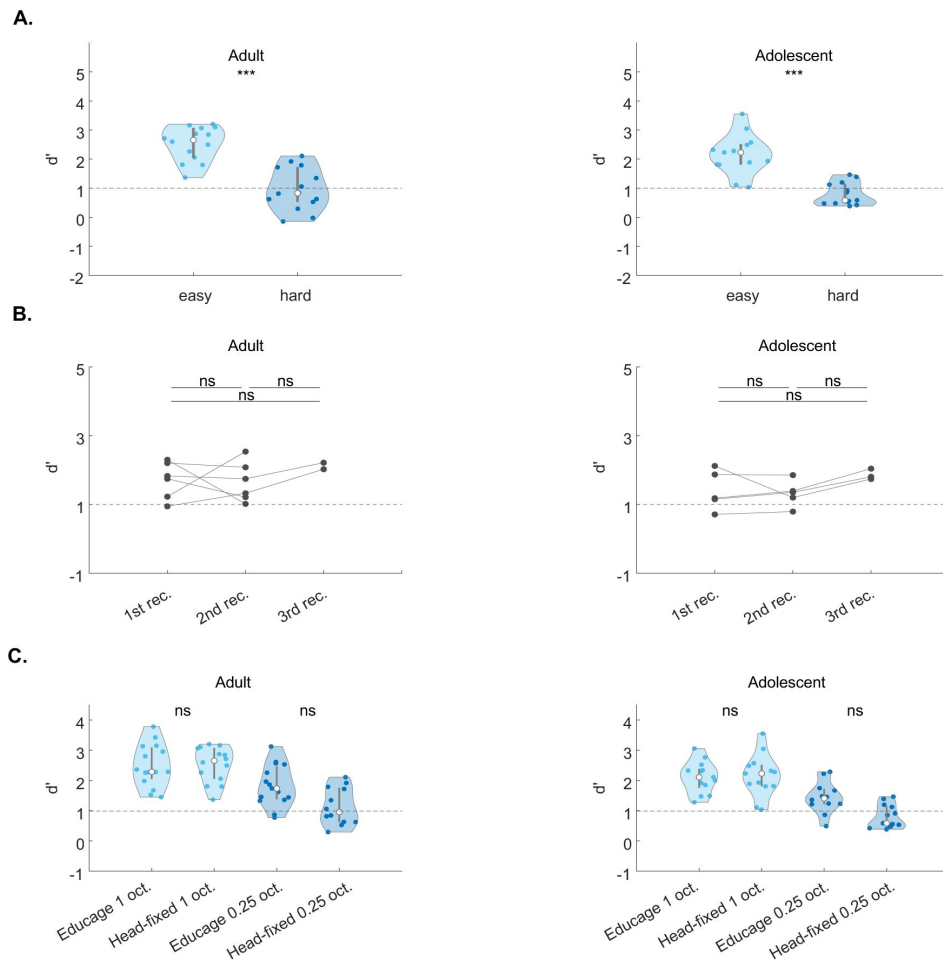
A. Experimental design for testing the role of ACx during tone discrimination in expert mice (adults only). **B.** Protocol for transient optogenetic suppression (light pulse duration was -50ms from tone onset, to +50ms from tone offset). Created with *BioRender.com*. **C.** Injection sites and optical fiber implantation for the experimental (GtACR2; top) and control groups (dTomato; bottom). **D.** Lick ratio for Go and No-Go stimuli under light-off conditions (grey) as compared to light-on conditions (red) in experimental (GtACR2, n = 13; left; Go stimuli: p = 0.0001; No-Go stimuli: p = 0.0107; one-sample Wilcoxon sign ranked test after Bonferroni correction) and control mice (dTomato, n = 8; right; Go stimuli: p = 0.4263; No-Go stimuli: p = 0.2953; one-sample Wilcoxon sign ranked test after Bonferroni correction). **E.** Lick ratio under light-off trials after light-on (red) or light off trials (grey) in experimental (GtACR2, n = 13; left; Go stimuli: p = 0.3864; No-Go stimuli: p = 0.2231; one-sample Wilcoxon sign ranked test after Bonferroni correction) and control (dTomato, n = 8; right; Go stimuli: p = 0.58341; No-Go stimuli: p = 0.3214; one-sample Wilcoxon sign ranked test after Bonferroni correction). **F.** Behavioral performance (d') per session under light-off conditions in easy (light blue) and hard (dark blue) task, as compared to light-on conditions (red) in experimental (GtACR2, n = 13; left; easy task: p = 0.0010, hard task: p = 0.0007, two-sample Wilcoxon rank sum test after Bonferroni correction; light-on p = 0.0010; light-off: p = 0.0398, one-sample Wilcoxon sign ranked test after Bonferroni correction) and control mice (dTomato, n = 8; right, easy task: p = 0.9999, hard task: p = 0.7422, two-sample Wilcoxon rank sum test after Bonferroni correction; light on p = 0.0312; light off: p = 0.0234, one-sample Wilcoxon sign ranked test after Bonferroni correction).



Supplementary Figure 3-3

Verification of GtACR2 expression.

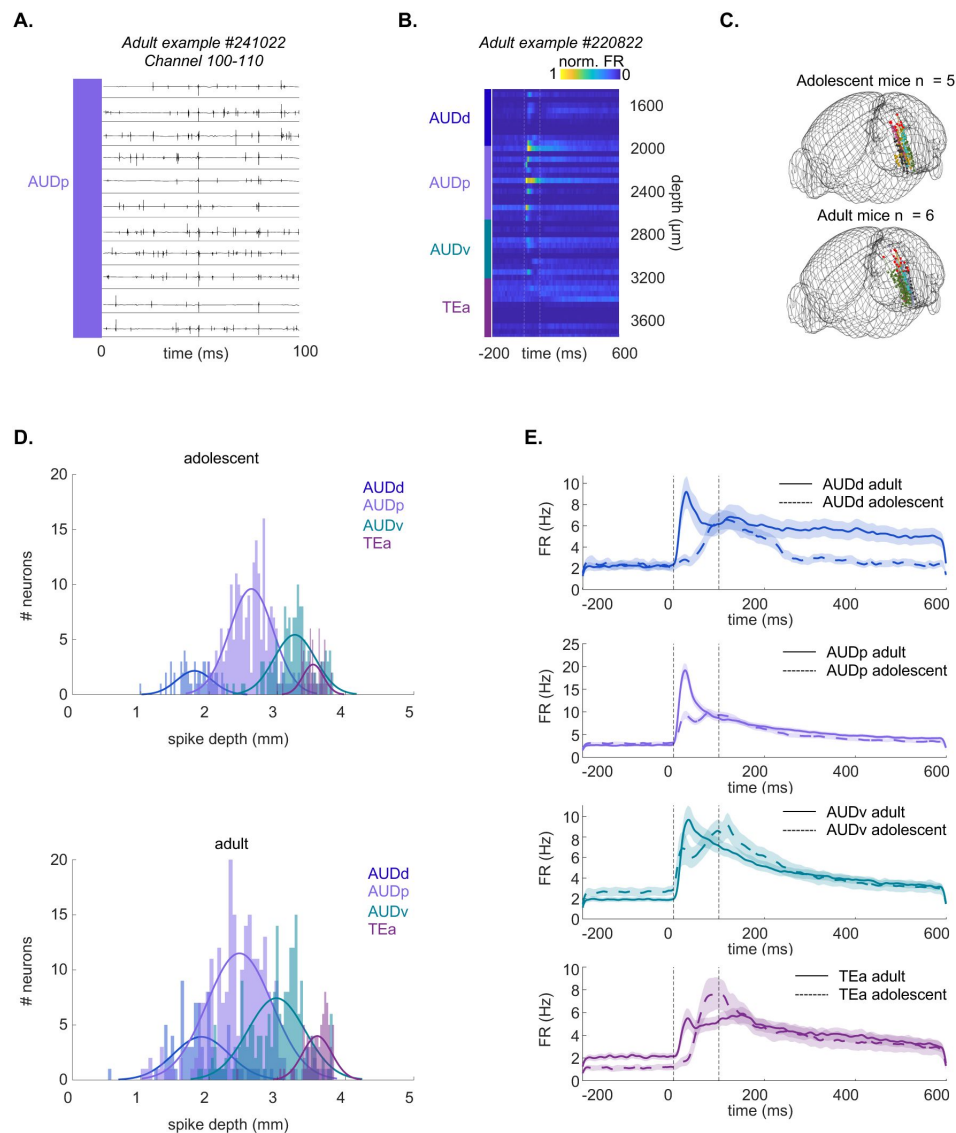
Photomicrograph (left), and reconstruction (right) of the GtACR2 infected area (red) per mouse. Reconstruction followed the coordinates of the Allen-CFF template-atlas. ACx is highlighted in white and GtACR2-expression in red.



Supplementary Figure 3-4

Adolescent and adult mice performed similarly throughout recordings as well as between the head-fixed configuration and the Educage.

A. Behavioral performance (d') in the easy task (light blue) and hard task (dark blue) for adolescent (recording = 13; left) and adult (recording = 14; right) recordings at the behavioral criterion of $d' > 1$ (adolescent mice: signed rank: 91, $p = 2.4414 \times 10^{-4}$; adult mice: signed rank: 105, $p = 1.2207 \times 10^{-4}$, Wilcoxon sign ranked test). Behavioral threshold of $d' = 1$ highlighted in the dashed line. **B.** Behavioral performance (average d' of the easy and the hard task) for every mouse per recording for adolescent mice ($n = 5$; left; 1st rec.: $p = 0.8125$; 2nd rec.: $p = 0.9999$; 3rd rec.: $p = 0.9999$, Wilcoxon sign ranked test, after Bonferroni correction), and adult mice ($n = 6$; right; 1st rec.: $p = 0.8438$; 2nd rec.: $p = 0.9999$; 3rd rec.: $p = 0.9999$, Wilcoxon sign ranked test, after Bonferroni correction). Behavioral threshold of $d' = 1$ highlighted in the dashed line. **C.** Comparison of behavioral performance in the head-fixed configuration and the Educage for adult mice (left; easy task: $p = 0.9960$; hard task: $p = 0.2159$, Kruskal Wallis test after Bonferroni correction), and adolescent mice (right; easy task: $p = 0.9973$; hard task: $p = 0.1505$, Kruskal Wallis test after Bonferroni correction) in the easy (light blue) and the hard (dark blue) task.



Supplementary Figure 4-1

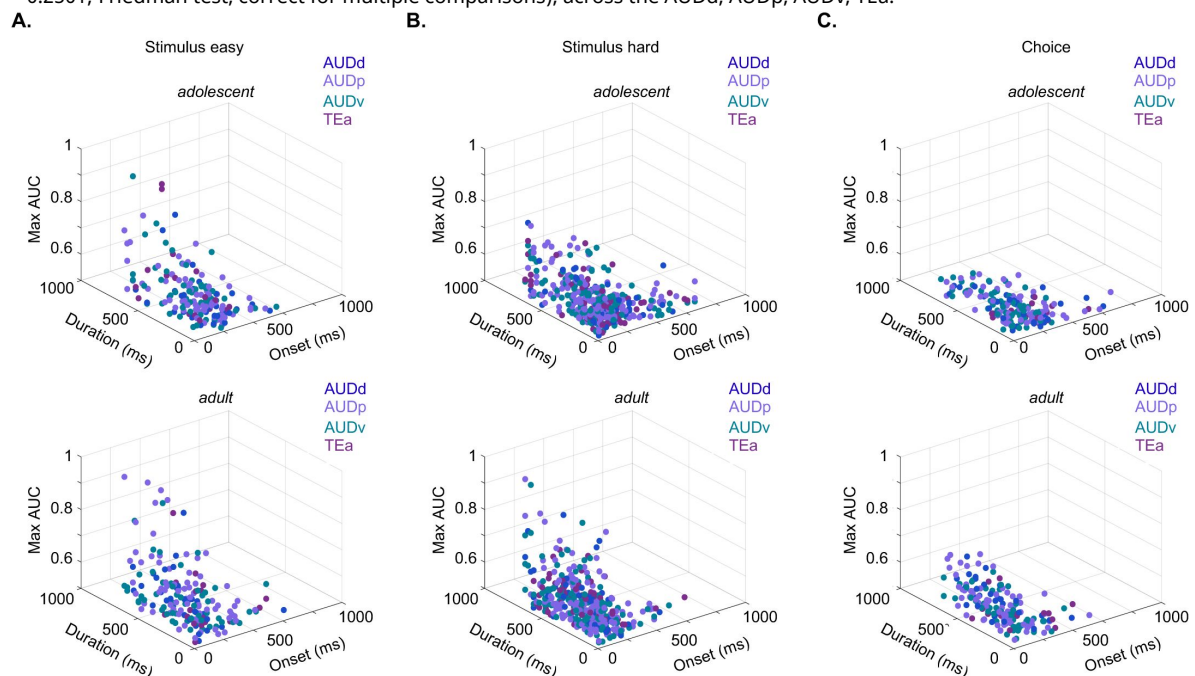
Probe reconstruction and activity profile across ACx regions.

A. Example voltage trace during a tone (100ms) across 11 channels. **B.** Example PSTH per depth (50 μ m bins) across AUDd, AUDp, AUDv and TEa (3850 maximal depth μ m). **C.** 3D-Reconstruction of recording sites in adolescent (n = 5; left) and adult (n = 6; right) mice. **D.** Spike-depth of excitatory tone-responsive L5/6 neurons in AUDd, AUDp, AUDv and TEa of adolescent (top) and adult (bottom) recordings. **E.** Population PSTH from -200ms to 600ms after tone onset in AUDd (blue), AUDp (purple), AUDv (magenta) and TEa (green) for adolescent neurons (dashed line) and adult neurons (solid line).

Supplementary Figure 4-2

The neuronal discriminability of stimulus- and choice related activity is similar across auditory sub-regions.

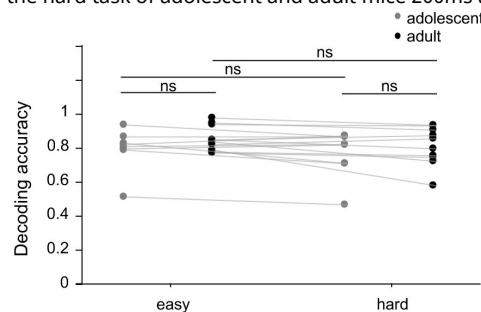
A. Onset-latency of discriminability (ms), duration of discriminability (ms), and maximal discriminability (AUC) of neurons that showed significant discriminability (exceeded 3 STD of the shuffled distribution) in the easy task (adolescent neurons = 178 (93%), mice = 4, recordings = 6; adult n = 346 (97%), mice = 4, recordings = 8; adolescent: onset-latency of discriminability: $p = 0.5310$, duration of discriminability: $p = 0.5418$, maximal discriminability: $p = 0.7212$; adult: onset-latency of discriminability: $p = 0.3810$, duration of discriminability: $p = 0.6105$, maximal discriminability: $p = 0.9115$; Friedman test, correct for multiple comparisons). **B.** Same as 'A', but in the hard task (adolescent neurons = 399 (93%), mice = 5, recordings = 10; adult n = 544 (97%), mice = 6, recordings = 12; adolescent: onset-latency of discriminability: $p = 0.9402$, duration of discriminability: $p = 0.3388$, maximal discriminability: $p = 0.6685$; adult: onset-latency of discriminability: $p = 0.2425$, duration of discriminability: $p = 0.5700$, maximal discriminability: $p = 0.1011$; Friedman test, correct for multiple comparisons). **C.** Same as 'A', but for choice-related activity (adolescent neurons = 181 (95%), mice = 4, recordings = 9; adult n = 339 (95%), mice = 4, recordings = 7; adolescent: onset-latency of discriminability: $p = 0.3975$, duration of discriminability: $p = 0.6823$, maximal discriminability: $p = 0.2866$; adult: onset-latency of discriminability: $p = 0.0881$, duration of discriminability: $p = 0.8185$, maximal discriminability: $p = 0.2501$; Friedman test, correct for multiple comparisons), across the AUDd, AUDp, TEa.

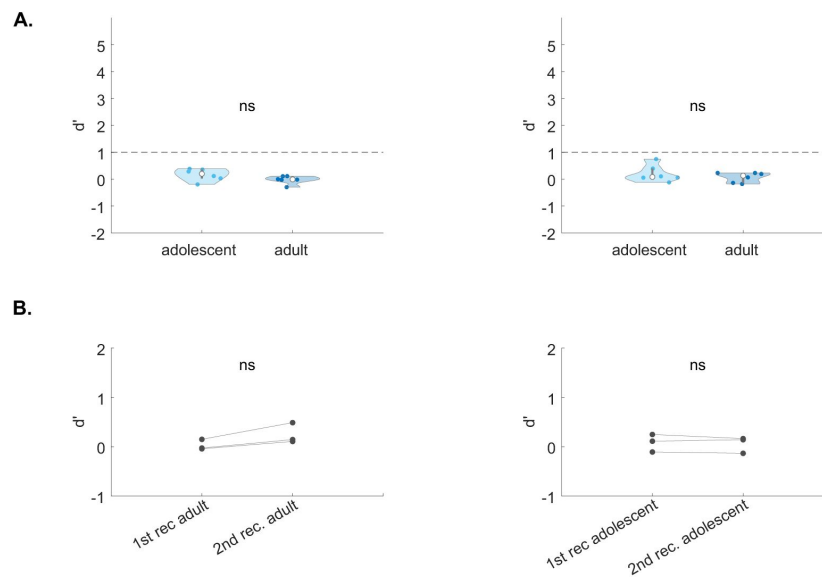


Supplementary Figure 5-1

Decoding accuracy of the first 200ms after the response window.

LDA decoding accuracy the easy and the hard task of adolescent and adult mice 200ms after the response window.

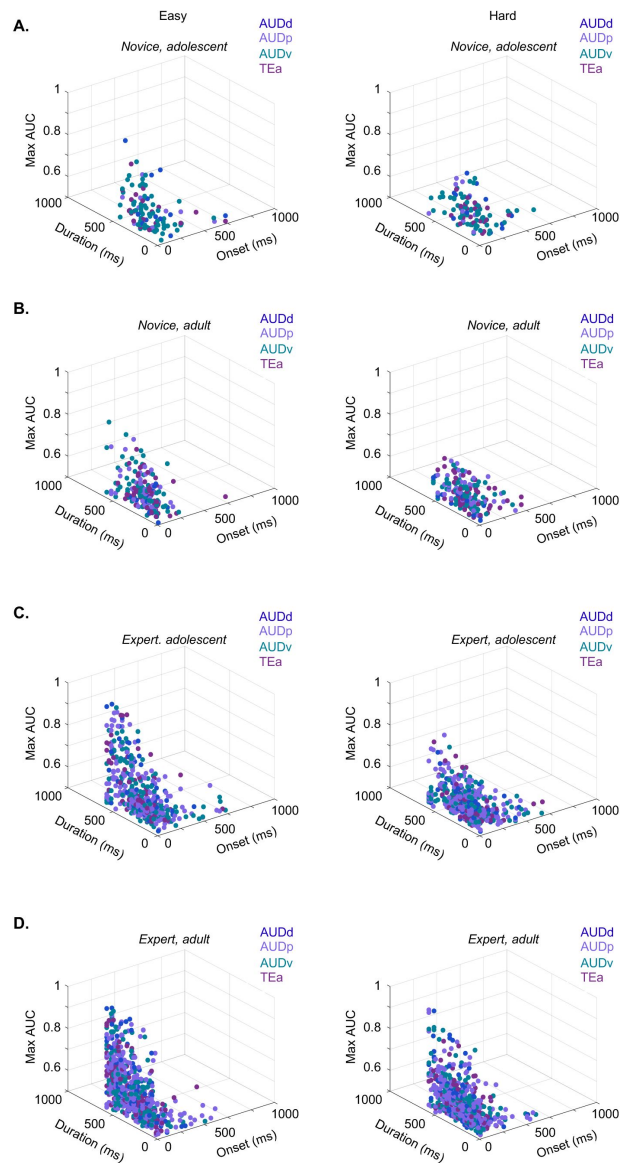




Supplementary Figure 6-1

Behavioral performance of novice mice.

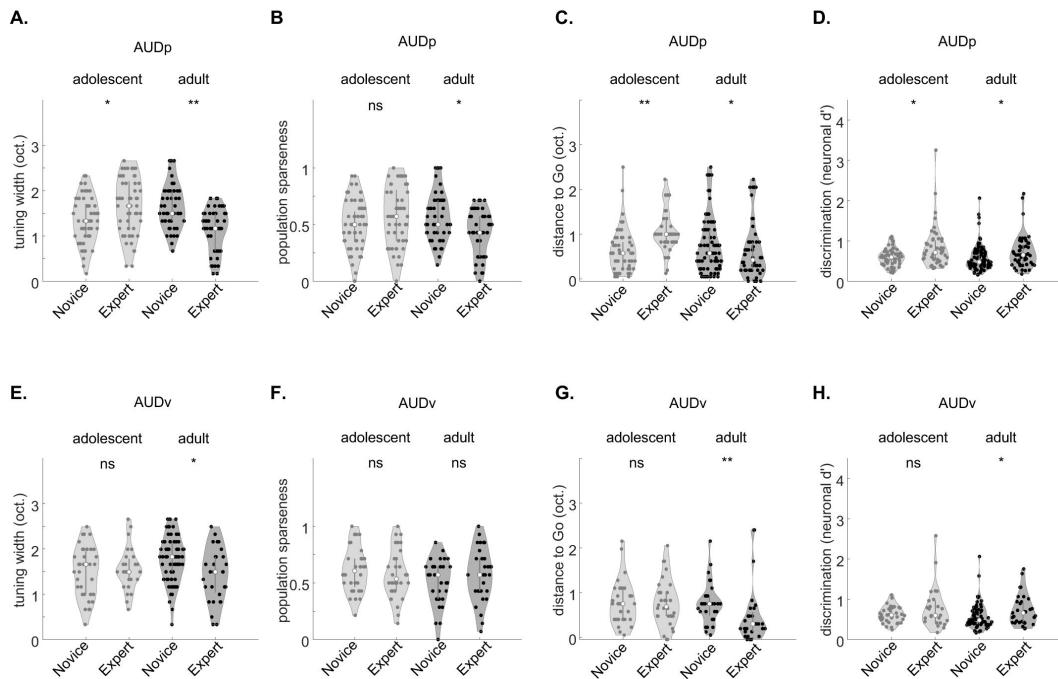
A. Behavioral performance (d') in the easy task (light blue) and hard task (dark blue) for adolescent ($n = 6$; left) and adult ($n = 6$; right) mice at the behavioral criterion of $d' > 1$ (adolescent mice: signed rank: 19, $p = 0.0938$; adult mice: signed rank: 15, $p = 0.4375$, Wilcoxon sign ranked test). Behavioral threshold of $d' = 1$ highlighted in the dashed line. **B.** Behavioral performance (average d' of the easy and the hard task) for every mouse per recording for adult mice ($n = 3$; left; signed rank: 0, $p = 0.2500$ Wilcoxon sign ranked test). adolescent mice ($n = 3$; right; signed rank: 0, $p = 0.2500$, Wilcoxon sign ranked test). Behavioral threshold of $d' = 1$ highlighted in the dashed line.



Supplementary Figure 6-2

The neuronal discriminability of Easy and Hard Go and No-Go are distributed similar across auditory sub-regions in adolescent and adult novice and expert mice.

Onset-latency of discriminability (ms), duration of discriminability (ms), and maximal discriminability (AUC) of neurons that showed significant discriminability (exceeded 3 STD of the shuffled distribution). Left: easy task (Novice, adolescent $n = 108$ (83%), mice = 3, recording = 6; onset-latency of discriminability: $p = 0.2422$, duration of discriminability: $p = 0.5639$, maximal discriminability: $p = 0.2062$. Novice, adult $n = 179$ (97%), mice = 3, recording = 6, onset-latency of discriminability: $p = 0.8335$, duration of discriminability: $p = 0.8013$, maximal discriminability: $p = 0.1900$. Expert, adolescent $n = 450$ (97%), mice = 5, recording = 13, onset-latency of discriminability: $p = 0.0918$, duration of discriminability: $p = 0.4020$, maximal discriminability: $p = 0.0698$. Expert, adult $n = 598$ (99%), mice = 6, recording = 14, onset-latency of discriminability: $p = 0.6807$, duration of discriminability: $p = 0.7223$, maximal discriminability: $p = 0.7557$; Friedman test, correct for multiple comparisons); Right: hard task (Novice, adolescent $n = 108$ (83%), mice = 3, recording = 6, onset-latency of discriminability: $p = 0.6294$, duration of discriminability: $p = 0.0693$, maximal discriminability: $p = 0.0858$. Novice, adult $n = 181$ (96%), mice = 3, recording = 6, onset-latency of discriminability: $p = 0.2180$, duration of discriminability: $p = 0.3388$, maximal discriminability: $p = 0.0648$. Expert, adolescent $n = 440$ (95%), mice = 5, recording = 13, onset-latency of discriminability: $p = 0.9911$, duration of discriminability: $p = 0.9939$, maximal discriminability: $p = 0.4058$. Expert, adult $n = 589$ (98%), mice = 6, recording = 14, onset-latency of discriminability: $p = 0.7141$, duration of discriminability: $p = 0.4084$, maximal discriminability: $p = 0.6365$; Friedman test, correct for multiple comparisons); across the AUDd, AUDp, AUDv, TEa.



Supplementary Figure 7-1

Learning related changes in neuronal tuning properties in primary and secondary auditory cortex.

A. Tuning bandwidth in AUDp at 62 dB SPL of neurons in adolescents and adults. Side by side comparisons of novice versus experts. (adolescents $p = 0.0438$, adults $p = 0.0001$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **B.** Same as 'A' for the population sparseness in AUDp (adolescents $p = 0.5724$, adults $p = 0.0066$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **C.** Same as 'A' for the distance (in octaves) between the best-frequency of each neuron to the easy Go-stimulus in AUDp (adolescents $p = 0.0001$, adults $p = 0.0201$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **D.** Same as 'A' for the average neuronal d' of pairs of frequencies limited to the learned frequency spectrum in AUDp (adolescents $p = 0.0471$, adults $p = 0.0321$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons). **E-H.** Same as A-D but in the AUDv. **E.** Adolescents $p = 0.9982$, adults $p = 0.0427$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons. **F.** Adolescents $p = 0.8841$, adults $p = 0.8031$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons. **G.** Adolescents $p = 0.9910$, adults $p = 0.0042$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons. **H.** Adolescents $p = 0.9438$, adults $p = 0.0108$, Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons.

Fixed Effects	Estimate	STE	T-Statistic	DF	P-Value	CI lower	CI upper
Intercept	5.441	0.2597	20.9483	1090	3.7563e-82	4.9314	5.9507
Lick Count	-3.4455	0.7075	-4.8697	1090	1.2827e-06	-4.8339	-2.0572
Lick Latency	-0.0086	0.0006	-15.0051	1090	2.0896e-46	-0.0098	-0.0075
d'	0.0584	0.1102	0.5303	1090	0.5960	-0.1577	0.2746
Count- Latency	0.0079	0.0026	3.0258	1090	0.0076	0.0028	0.013
Count - d'	3.5431	1.1445	3.0956	1090	0.006	1.2973	5.7889
Latency – d'	-0.0004	0.0004	-0.9635	1090	0.9999	-0.0011	0.0004
Latency – Count - d'	-0.0091	0.0041	-2.2524	1090	0.0735	-0.0171	-0.0012

Supplementary Table 1.

Adolescent and adult mice exhibit different lick behavior in the task.

Linear mixed-effects models of the fixed effects of lick count (until reward or punishment delay), lick latency, cumulative discriminability (d') (including the interaction effects of lick count and lick latency, lick count and d' , lick latency and d' , and lick latency, lick count and d') during the minimal number of trials shared between all mice (148 trials; Number of observations = 1098, Fixed effects coefficients = 8, Random effects coefficients = 14, Covariance parameters = 3). Coefficient estimates, STE, T-statistic, degrees of freedom, p-values (adjusted for post-hoc multiple comparisons with Bonferroni-method), lower and higher CI are listed in the table. The model includes random effects coefficients per mouse (11 mice in total) and 3 recordings per mouse (see *methods, equation 8*). Model structure: Lick Count ~ Group * Lick Latency * d' prime + (1 | Mouse ID) + (1 | Recording ID).

Expert	adolescent			adult	
	areas	total	excited	total	excited
	AUDd	177	56 (32%)	154	82 (53%)
	AUDp	442	210 (48%)	535	245 (46%)
	AUDv	338	132 (39%)	343	197 (57%)
	TEa	188	65 (35%)	235	75 (32%)
Novice	adolescent			adult	
	areas	total.	excited	total	excited
	AUDd	89	22 (25%)	96	17 (18%)
	AUDp	167	31 (19%)	181	49 (27%)
	AUDv	302	47 (16%)	180	65 (36%)
	TEa	99	30 (33%)	146	55 (38%)

Supplementary Table 2

Neuronal statistics in expert and novice recordings during task engagement.

Acquired single units, acquired tone-excited units (percentage of tone-excited units relative to total units) in the AUDd, AUDp, AUDv, and TEa of adolescent and adult mice in experts (top) and novice (bottom).

AUDd	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	2.4274	0.4188	2.2271	0.3668	0.1414	-0.0111	0.2997	0.9801
evoked FR	14.5588	1.6046	22.1731	2.9083	-0.306	-0.3798	-0.1396	0.1537
FR coeff. var	2.4484	0.2496	1.7209	0.1238	0.4794	0.0573	0.3294	0.0146
latency to peak	134.5179	8.7109	132.061	16.1264	0.1266	0.1826	0.5423	0.0006
FWHM	275.7321	26.1309	245.4634	24.4123	0.1317	0.0409	0.2443	0.2642
min. latency	73.5638	3.3501	48.9258	3.1867	0.7905	0.3529	0.6238	3.6051E-06
% trials resp.	0.3809	0.0364	0.4815	0.0328	-0.2993	-0.3032	-0.076	0.0422
lifetime sparse.	0.1744	0.0291	0.1506	0.0183	0.0401	-0.3136	-0.0548	0.7897
AUDp	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	3.158	0.2654	2.7504	0.2546	0.1752	-0.0103	0.2951	0.0115
evoked FR	24.5512	1.6812	39.4108	2.6007	-0.432	-0.4069	-0.1391	5.7054E-05
FR coeff. var	1.6895	0.0872	1.4396	0.084	0.2913	0.0522	0.3326	0.0004
latency to peak	102.981	7.9174	50.4	3.0841	0.7259	0.1981	0.5378	1.0347E-11
FWHM	241.2571	13.7176	150.2408	12.0679	0.4569	0.0357	0.2445	8.7880E-10
min. latency	47.4148	1.8017	31.4164	1.5054	0.6936	0.3569	0.6333	5.605E-12
% trials resp.	0.517	0.0208	0.6	0.0209	-0.2354	-0.2853	-0.0645	0.0026
lifetime sparse.	0.1459	0.0115	0.1875	0.0118	-0.2739	-0.3028	-0.0485	0.0049
AUDv	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	2.627	0.4173	1.9547	0.2387	0.2687	-0.0121	0.2867	0.1084
evoked FR	21.8639	2.8512	24.4575	1.9949	-0.1752	-0.3839	-0.1324	0.4306
FR coeff. var	1.8877	0.1089	1.7797	0.1009	0.1175	0.0477	0.3122	0.0927
latency to peak	126.1818	9.6388	135.5939	10.1844	-0.0138	0.1951	0.5405	0.2089
FWHM	226.5985	16.5937	241.7817	15.1074	-0.0613	0.0334	0.2419	0.7051
min. latency	59.9072	2.3282	51.9611	2.1164	0.2969	0.3359	0.6227	0.0062
% trials resp.	0.4461	0.0253	0.5074	0.023	-0.1811	-0.2962	-0.0823	0.1396
lifetime sparse.	0.1593	0.0153	0.2033	0.0127	-0.3125	-0.3089	-0.0527	0.0038
TEa	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	1.2426	0.2315	2.1943	0.2731	-0.5523	-0.0096	0.303	0.0059
evoked FR	17.567	2.5111	16.5377	1.7923	-0.0353	-0.4066	-0.1551	0.8271
FR coeff. var	1.9063	0.1606	2.1603	0.2278	-0.0467	0.0585	0.3122	0.9088
latency to peak	170.0645	18.1797	145.8133	15.4745	0.2663	0.2071	0.5479	0.2447
FWHM	259.6935	25.6747	297.2267	23.7387	-0.1590	0.038	0.2512	0.4505
min. latency	73.5202	2.8156	59.9417	2.8259	0.5712	0.3687	0.6166	0.0013
% trials resp.	0.4412	0.0375	0.4515	0.0347	-0.0491	-0.3016	-0.0768	0.909
lifetime sparse.	0.2156	0.0238	0.1819	0.0206	0.2167	-0.3118	-0.0521	0.1658

Supplementary Table 3

Adolescent and adult expert mice have distinct firing properties in different sub-regions.

Mean and standard error, mean effect size (robust Cohen's D), lower and upper Confidence Interval (CI) and p-value (Wilcoxon rank-sum test, adjusted for multiple comparisons with Bonferroni method) of the average baseline FR (Hz), evoked FR (Hz), coefficient of variance of FR, latency to peak of maximal FR (ms), full-width-half maximum of peak FR (ms), minimal latency of first spike (ms), fraction of responsive trials, lifetime sparseness of all adolescent and adult neurons from tone-onset to 50ms after tone offset across all stimuli in AUDd, AUDp, and AUDv, and TEa (significant p-values are highlighted in bold).

adolescent	P-Value**					
neuronal property	AUDd - AUDp	AUDd - AUDv	AUDd - TEa	AUDp - AUDv	AUDp - TEa	AUDv - TEa
spontaneous FR	0.3379	0.9998	0.3018	0.0793	0.0004	0.189
evoked FR	0.0276	0.5874	0.9983	0.2043	0.0367	0.6856
FR coeff. var	0.0078	0.3519	0.5517	0.2063	0.3573	0.9992
latency to peak	0.0001	0.0878	0.9894	0.0001	0.0001	0.1681
FWHM	0.7022	0.4202	0.998	0.8836	0.8076	0.5188
min. latency	0.0001	0.0263	0.9819	0.0002	0.0001	0.0045
% trials resp.	0.014	0.4845	0.6757	0.1778	0.3250	0.9992
lifetime sparse.	0.9998	0.9453	0.0255	0.8245	0.0016	0.027
adult	P-Value**					
neuronal property	AUDd - AUDp	AUDd - AUDv	AUDd - TEa	AUDp - AUDv	AUDp - TEa	AUDv - TEa
spontaneous FR	0.999	0.4036	0.9081	0.2066	0.7793	0.0985
evoked FR	0.0004	0.9843	0.6787	0.0001	0.0001	0.3436
FR coeff. var	0.0119	0.9605	0.7967	0.0038	0.0003	0.4040
latency to peak	0.0001	0.6264	0.0672	0.0001	0.0001	0.3092
FWHM	0.0023	0.9068	0.1614	0.0001	0.0001	0.2782
min. latency	0.0001	0.8888	0.0409	0.0001	0.0001	0.0744
% trials resp.	0.0192	0.9373	0.9163	0.0114	0.0018	0.5453
lifetime sparse.	0.2722	0.0203	0.3678	0.4004	0.9963	0.8102

** Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons

Supplementary Table 4

Adolescent and adult firing properties of expert mice are distinct between different sub-regions.

P-values of Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons of the average baseline FR (Hz), evoked FR (Hz), coefficient of variance of FR, latency to peak of maximal FR (ms), full-width-half maximum of peak FR (ms), minimal latency of first spike (ms), fraction of responsive trials, lifetime sparseness of all adolescent and adult neurons from tone-onset to 50ms after tone offset across all stimuli AUDd – AUDp, AUDd – AUDv, AUDd – TEa, AUDp – AUDv, AUDp – Tea, and AUDv –Tea (significant p-values are highlighted in bold).

AUDd	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	2.4274	0.4188	2.2271	0.3668	0.3224	-1.0907	1.9771	0.9646
evoked FR	14.5588	1.6046	22.1731	2.9083	0.323	-0.4649	1.1014	0.4924
FR coeff. var	2.4484	0.2496	1.7209	0.1238	-0.2152	-0.9768	0.6759	0.3410
latency to peak	134.5179	8.7109	132.061	16.1264	-0.0601	-1.2314	0.773	0.7230
FWHM	275.7321	26.1309	245.4634	24.4123	-0.0333	-0.7347	0.701	0.3998
min. latency	73.5638	3.3501	48.9258	3.1867	-0.5259	-1.2905	0.1873	0.1265
% trials resp.	0.3809	0.0364	0.4815	0.0328	0.1885	-0.5429	1.2357	0.7397
lifetime sparse.	0.1744	0.0291	0.1506	0.0183	0.5616	-0.0983	1.7234	0.3410
AUDp	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	3.158	0.2654	2.7504	0.2546	0.9234	0.5075	1.4603	1.41592E-05
evoked FR	24.5512	1.6812	39.4108	2.6007	0.8286	0.4288	1.4088	0.0015
FR coeff. var	1.6895	0.0872	1.4396	0.084	-0.6202	-1.0043	-0.2827	0.0012
latency to peak	102.981	7.9174	50.4	3.0841	-0.5161	-1.0001	-0.0189	0.0403
FWHM	161.2571	13.7176	150.2408	12.0679	-0.1835	-0.5979	0.393	0.0495
min. latency	47.4148	1.8017	31.4164	1.5054	-0.4746	-0.9192	-0.1092	0.0406
% trials resp.	0.517	0.0208	0.6	0.0209	0.667	0.2199	1.1501	0.0007
lifetime sparse.	0.1459	0.0115	0.1875	0.0118	0.3354	-0.2348	0.8616	0.1258
AUDv	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	2.627	0.4173	1.9547	0.2387	0.1787	-0.2081	0.6092	0.0963
evoked FR	21.8639	2.8512	24.4575	1.9949	-0.5624	-0.929	-0.2553	0.0186
FR coeff. var	1.8877	0.1089	1.7797	0.1009	0.1859	-0.2146	0.6288	0.1282
latency to peak	126.1818	9.6388	135.5939	10.1844	0.7395	0.3299	1.1785	0.0001
FWHM	226.5985	16.5937	241.7817	15.1074	0.518	0.1243	1.1608	0.0002
min. latency	59.9072	2.3282	51.9611	2.1164	0.8709	0.4572	1.4074	5.4742E-06
% trials resp.	0.4461	0.0253	0.5074	0.023	-0.2375	-0.5761	0.0878	0.1676
lifetime sparse.	0.1593	0.0153	0.2033	0.0127	0.3121	-0.0957	0.6979	0.0608
TEa	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	1.2426	0.2315	2.1943	0.2731	0.8305	0.1538	2.1521	0.0082
evoked FR	17.567	2.5111	16.5377	1.7923	-0.6609	-0.9696	-0.2934	0.0108
FR coeff. var	1.9063	0.1606	2.1603	0.2278	-0.0567	-0.457	0.6441	0.8526
latency to peak	170.0645	18.1797	145.8133	15.4745	0.3679	-0.0907	1.0797	0.0440
FWHM	299.6935	25.6747	277.2267	23.7387	1.707	1.0083	3.2231	8.99288E-06
min. latency	73.5202	2.8156	59.9417	2.8259	0.5511	0.002	1.1864	0.0112
% trials resp.	0.4412	0.0375	0.4515	0.0347	-0.0406	-0.4834	0.4156	0.8385
lifetime sparse.	0.2156	0.0238	0.1819	0.0206	0.3157	-0.3936	0.9299	0.7058

Supplementary Table 5

Novice, adolescent and adult mice have distinct firing properties in different sub-regions.

Mean and standard error (STE), mean effect size (robust Cohen's D), lower and upper Confidence Interval (CI) and p-value (Wilcoxon rank-sum test, adjusted for multiple comparisons with Bonferroni method) of the average baseline FR (Hz), evoked FR (Hz), coefficient of variance of FR, latency to peak of maximal FR (ms), full-width-half maximum of peak FR (ms), minimal latency of first spike (ms), fraction of responsive trials, lifetime sparseness of all adolescent and adult neurons from tone-onset to 50ms after tone offset across all stimuli in AUDd, AUDp, and AUDv, and TEa (significant p-values are highlighted in bold).

adolescent	P-Value**					
neuronal property	AUDd - AUDp	AUDd - AUDv	AUDd - TEa	AUDp - AUDv	AUDp - TEa	AUDv - TEa
spontaneous FR	0.3238	0.5109	0.1699	0.9421	0.9336	0.6709
evoked FR	0.6254	0.9512	0.9985	0.6134	0.0418	0.8233
FR coeff. var	0.2319	0.8251	0.3252	0.3928	0.0998	0.5767
latency to peak	0.188	0.8251	0.9998	0.0002	0.0449	0.7502
FWHM	0.7308	0.3703	0.2985	0.0009	0.0032	0.9723
min. latency	0.3989	0.7007	0.7987	0.0008	0.0137	0.9999
% trials resp.	0.2022	0.7372	0.2238	0.4637	0.0996	0.5149
lifetime sparse.	0.9927	0.9657	0.8218	0.9944	0.0649	0.9267
adult	P-Value**					
neuronal property	AUDd - AUDp	AUDd - AUDv	AUDd - TEa	AUDp - AUDv	AUDp - TEa	AUDv - TEa
spontaneous FR	0.4679	0.9932	0.993	0.0533	0.3267	0.8491
evoked FR	0.9866	0.263	0.0141	0.0096	0.0001	0.2526
FR coeff. var	0.9344	0.0143	0.0047	0.0032	0.0006	0.9345
latency to peak	0.8167	0.1522	0.1139	0.3118	0.0279	0.9939
FWHM	0.3574	0.2036	0.0662	0.9806	0.0413	0.848
min. latency	0.16	0.0009	0.0068	0.0884	0.0392	0.8875
% trials resp.	0.9998	0.0798	0.0174	0.0021	0.0001	0.804
lifetime sparse.	0.9933	0.9319	0.9764	0.9665	0.9972	0.9932
** Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons						

Supplementary Table 6

Adolescent and adult firing properties of novice mice are distinct between different sub-regions.

P-values of Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons of the average baseline FR (Hz), evoked FR (Hz), coefficient of variance of FR, latency to peak of maximal FR (ms), full-width-half maximum of peak FR (ms), minimal latency of first spike (ms), fraction of responsive trials, lifetime sparseness of all adolescent and adult neurons from tone-onset to 50ms after tone offset across all stimuli AUDd - AUDp, AUDd - AUDv, AUDd - TEa, AUDp - AUDv, AUDp - TEa, and AUDv -Tea (significant p-values are highlighted in bold).

Group					
Expert	adolescent			adult	
	areas	total	excited	total	excited
	AUDd	29	1 (0.03%)	64	13 (0.2%)
	AUDp	147	44 (0.30%)	111	27 (0.24%)
	AUDv	137	35 (0.26%)	109	25 (0.23%)
	TEa	35	0 (0%)	124	19 (0.15%)
Novice	adolescent			adult	
	areas	total	Excited	total	Excited
	AUDd	89	6 (0.07%)	96	8 (0.08%)
	AUDp	167	25 (0.15%)	181	24 (0.13%)
	AUDv	302	44 (0.15%)	180	45 (0.25%)
	TEa	99	17 (0.17%)	146	46 (0.32%)

Supplementary Table 7

Neuronal statistics during passive FRA protocol in expert and novice mice.

Acquired single units, acquired tone-modulated units, and percentage of modulated units to all acquired units in the AUDd, AUDp, AUDv, and TEa of adolescent and adult mice during passive-listening recordings.

AUDp	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	15.0108	2.552	16.1657	2.0109	-0.1182	-0.6373	0.3627	0.8636
evoked FR	33.9654	8.0969	40.81	4.012	-0.0807	-0.5791	0.4458	0.7709
FR coeff. var	0.847	0.1456	0.6278	0.1104	0.5127	0.0149	1.0775	0.0646
latency to peak	110.8182	8.3515	61.963	17.0012	0.4648	0.0048	0.9211	0.0472
FWHM	178.3409	21.0212	68.5185	26.5098	0.6418	0.2081	1.0021	0.0032
min. latency	55.5352	6.6073	48.058	4.6893	0.242	-0.2302	0.7889	0.0389
% trials resp.	0.6335	0.0827	0.5972	0.5320	0.0716	-0.473	0.507	0.9951
lifetime sparse.	0.4198	0.0458	0.3383	0.3580	0.2868	-0.2248	0.7884	0.1048
AUDv	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	18.9295	1.8448	8.7466	3.8163	0.47	-0.0115	0.8816	0.0174
evoked FR	46.1598	5.3813	21.9135	8.8535	0.6166	0.0396	1.2171	0.0068
FR coeff. var	1.0532	0.232	0.9158	0.1627	0.2143	-0.3916	0.9365	0.1785
latency to peak	109.4857	19.6319	89.68	12.9514	0.3197	-0.3407	0.8796	0.18106
FWHM	146.0857	30.3333	119.2	26.5399	0.1272	-0.4042	0.7038	0.0439
min. latency	60.4915	8.3884	49.5169	4.3829	0.2985	-0.2216	0.8729	0.0433
% trials resp.	0.6946	0.0804	0.3925	0.0545	0.7253	0.1905	1.5257	0.0027
lifetime sparse.	0.3773	0.0533	0.2565	0.0367	0.4362	-0.0431	1.1493	0.0451

Supplementary Table 8

Adolescent and adult mice have distinct firing properties across different sub-regions of ACx—passive listening.

Mean and standard error, mean effect size (robust Cohen's D), lower and upper Confidence Interval (CI) and p-value (Wilcoxon rank-sum test) of the average baseline FR (Hz) (AUDp vs. AUDv: adolescent p = 0.8551; adult p = 0.9711), evoked FR (Hz) (AUDp vs. AUDv: adolescent p = 0.4125; adult p = 0.9954), coefficient of variance of FR (AUDp vs. AUDv: adolescent p = 0.4354; adult p = 0.8800), latency to peak of maximal FR (ms) (AUDp vs. AUDv: adolescent p = 0.5871; adult p = 0.9985), full-width-half maximum of peak FR (ms) (AUDp vs. AUDv: adolescent p = 0.7223; adult p = 0.4628), minimal latency of first spike (ms) (AUDp vs. AUDv: adolescent p = 0.5936; adult p = 0.5669), fraction of responsive trials (AUDp vs. AUDv: adolescent p = 0.3838; adult p = 0.9924), lifetime sparseness (AUDp vs. AUDv: adolescent p = 0.3792; adult p = 0.9341) of all adolescent and adult neurons from tone-onset to 50ms after tone offset across all stimuli in AUDp, and AUDv (significant p-values are highlighted in bold).

Group	Expert			Novice		
	Mice	Recording	Neurons	Mice	Recording	Neurons
Figure 1	15 (adolescent) 15 (adult)	none	none	none	none	none
Figure 2	15 (adolescent) 15 (adult)	none	none	none	none	none
Figure 3	5 (adolescent) 6 (adult)	none	none	none	none	none
Figure S3-2	4 (GtACR2) 3 (dTomato)	none	none	none	none	none
Figure 4	5 (adolescent) 6 (adult)	13 (adolescent) 14 (adult)	1145 (adolescent) 1267 (adult)	none	none	none
Figure 5	5 (adolescent) 6 (adult)	13 (adolescent) 14 (adult)	1145 (adolescent) 1267 (adult)	none	none	none
Figure 6	5 (adolescent) 6 (adult)	13 (adolescent) 14 (adult)	1145 (adolescent) 1267 (adult)	3 (adolescent) 3 (adult)	6 (adolescent) 6 (adult)	657 (adolescent) 603 (adult)
Figure 7	4 (adolescent) 4 (adult)	4 (adolescent) 4 (adult)	348 (adolescent) 408 (adult)	3 (adolescent) 3 (adult)	6 (adolescent) 6 (adult)	557 (adolescent) 503 (adult)

Supplementary Table 9

Overview of datasets analyzed.

Number mice, recordings, and neurons per learning stage (left: expert, right: novice) for every figure.

Data availability

The data that support the finding of this study are available through Zenodo (<https://zenodo.org/uploads/13933351>). Additional data is available from authors upon reasonable request.

Acknowledgements

We thank the Mizrahi lab for comments on the manuscript and Linda Willbrecht, and Madeleine Klinger for discussions. We thank Ido Maor, Or Yudco, Omri Gilday, and Meirav Givon for technical assistance. We also thank Ofer Yizhar for sharing reagents and Maya Groysman for preparing the viruses. This work was supported by an NSF-BSF grant to A.M. and S.D. (#2021776), stipends from the Minerva Foundation and the Israeli Ministry of Aliya and Integration to B.P., and the Gatsby Charitable Foundation. Adi Mizrahi is the Eric Roland Chair in Brain Sciences. **Figure 1a**, **3a**, S3-2b, 4a, 6a, 7a were created with biorender.com.

Additional information

Author Contributions

B.P., and A.M. designed the experiments. B.P, A.D., and A.L., performed the experiments, B.P, F.C., S.D., analyzed the data. B.P., and A.M. wrote the paper with comments from F.C. and S.D.

Inclusion and Diversity

We support inclusive, diverse, and equitable conduct of research

Code availability

The code is available through GitHub (<https://github.com/benne1295/Age-and-Learning-Shapes-Sound-Representations-in-Auditory-Cortex-During-Adolescence>).

Funding

United States-Israel Binational Science Foundation (2021776)

References

1. Hoskins D (2014) **Consequences of Parenting on Adolescent Outcomes** *Societies* **4**:506–31 <https://doi.org/10.3390/soc4030506> | Google Scholar
2. Hawkins JD, Catalano RF, Miller JY (1992) **Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention** *Psychol Bull* **112**:64–105 <https://doi.org/10.1037/0033-2909.112.1.64> | PubMed | Google Scholar
3. Fuhrmann D, Knoll LJ, Blakemore SJ (2015) **Adolescence as a Sensitive Period of Brain Development** *Trends Cogn Sci* **19**:558–66 <https://doi.org/10.1016/j.tics.2015.07.008> | PubMed | Google Scholar
4. Caras ML, Sanes DH (2019) **Neural Variability Limits Adolescent Skill Learning** *J Neurosci* **39**:2889–902 <https://doi.org/10.1523/JNEUROSCI.2878-18.2019> | PubMed | Google Scholar
5. Huyck JJ, Wright BA (2011) **Late maturation of auditory perceptual learning** *Dev Sci* **14**:614–21 <https://doi.org/10.1111/j.1467-7687.2010.01009.x> | PubMed | Google Scholar
6. Johnson C, Wilbrecht L (2011) **Juvenile mice show greater flexibility in multiple choice reversal learning than adults** *Dev Cogn Neurosci* **1**:540–51 <https://doi.org/10.1016/j.dcn.2011.05.008> | PubMed | Google Scholar
7. Romer D (2010) **Adolescent risk taking, impulsivity, and brain development: implications for prevention** *Dev Psychobiol* **52**:263–76 <https://doi.org/10.1002/dev.20442> | PubMed | Google Scholar
8. Wilbrecht L, Davidow JY (2024) **Goal-directed learning in adolescence: neurocognitive development and contextual influences** *Nat Rev Neurosci* **25**:176–94 <https://doi.org/10.1038/s41583-023-00783-w> | PubMed | Google Scholar
9. Hensch TK (2005) **Critical period plasticity in local cortical circuits** *Nat Rev Neurosci* **6**:877–88 <https://doi.org/10.1038/nrn1787> | PubMed | Google Scholar
10. Hensch TK. (2004) **Critical period regulation** *Annu Rev Neurosci* **27**:549–79 <https://doi.org/10.1146/annurev.neuro.27.070203.144327> | PubMed | Google Scholar
11. Reh RK, Dias BG, Nelson 3rd CA, et al. (2020) **Critical period regulation across multiple timescales** *Proc Natl Acad Sci U S A* **117**:23242–51 <https://doi.org/10.1073/pnas.1820836117> | PubMed | Google Scholar
12. Kral A (2013) **Auditory critical periods: a review from system’s perspective** *Neuroscience* **247**:117–33 <https://doi.org/10.1016/j.neuroscience.2013.05.021> | PubMed | Google Scholar
13. Barkat TR, Polley DB, Hensch TK (2011) **A critical period for auditory thalamocortical connectivity** *Nat Neurosci* **14**:1189–94 <https://doi.org/10.1038/nn.2882> | PubMed | Google Scholar

14. Sun YJ, Wu GK, Liu BH, Li P, Zhou M, Xiao Z, et al. (2010) **Fine-tuning of pre-balanced excitation and inhibition during auditory cortical development** *Nature* **465**:927–31 <https://doi.org/10.1038/nature09079> | PubMed | Google Scholar
15. Zhang LI, Bao S, Merzenich MM (2001) **Persistent and specific influences of early acoustic environments on primary auditory cortex** *Nat Neurosci* **4**:1123–30 <https://doi.org/10.1038/nn745> | PubMed | Google Scholar
16. Bhumika S, Nakamura M, Valerio P, Solyga M, Linden H, Barkat TR (2020) **A Late Critical Period for Frequency Modulated Sweeps in the Mouse Auditory System** *Cereb Cortex* **30**:2586–99 <https://doi.org/10.1093/cercor/bhz262> | PubMed | Google Scholar
17. Nakamura M, Valerio P, Bhumika S, Barkat TR (2020) **Sequential Organization of Critical Periods in the Mouse Auditory System** *Cell Rep* **32**:108070 <https://doi.org/10.1016/j.celrep.2020.108070> | PubMed | Google Scholar
18. Maor I, Shwartz-Ziv R, Feigin L, Elyada Y, Sompolinsky H, Mizrahi A (2019) **Neural Correlates of Learning Pure Tones or Natural Sounds in the Auditory Cortex** *Front Neural Circuits* **13** <https://doi.org/10.3389/fncir.2019.00082> | PubMed | Google Scholar
19. O'Sullivan C, Weible AP, Wehr M (2019) **Auditory Cortex Contributes to Discrimination of Pure Tones** *eNeuro* **6** <https://doi.org/10.1523/ENEURO.0340-19.2019> | PubMed | Google Scholar
20. Ceballo S, Piwkowska Z, Bourg J, Daret A, Bathellier B (2019) **Targeted cortical manipulation of auditory perception** *Neuron* **104**:1168–79 [Google Scholar](https://doi.org/10.1016/j.neuron.2019.06.003)
21. Kanold PO, Nelken I, Polley DB (2014) **Local versus global scales of organization in auditory cortex** *Trends Neurosci* **37**:502–10 <https://doi.org/10.1016/j.tins.2014.06.003> | PubMed | Google Scholar
22. Feigin L, Tasaka G, Maor I, Mizrahi A (2021) **Sparse Coding in Temporal Association Cortex Improves Complex Sound Discriminability** *J Neurosci* **41**:7048–64 <https://doi.org/10.1523/JNEUROSCI.3167-20.2021> | PubMed | Google Scholar
23. Bao S, Chang EF, Teng CL, Heiser MA, Merzenich MM (2013) **Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment** *Neuroscience* **248**:30–42 <https://doi.org/10.1016/j.neuroscience.2013.05.056> | PubMed | Google Scholar
24. Bao S, Chang EF, Woods J, Merzenich MM (2004) **Temporal plasticity in the primary auditory cortex induced by operant perceptual learning** *Nat Neurosci* **7**:974–81 <https://doi.org/10.1038/nn1293> | PubMed | Google Scholar
25. Han YK, Kover H, Insanally MN, Semerdjian JH, Bao S (2007) **Early experience impairs perceptual discrimination** *Nat Neurosci* **10**:1191–7 <https://doi.org/10.1038/nn1941> | PubMed | Google Scholar
26. Anbuhl KL, Yao JD, Hotz RA, Mowery TM, Sanes DH (2022) **Auditory processing remains sensitive to environmental experience during adolescence in a rodent model** *Nat Commun* **13**:2872 <https://doi.org/10.1038/s41467-022-30455-9> | PubMed | Google Scholar
27. Jaramillo S, Zador AM (2011) **The auditory cortex mediates the perceptual effects of acoustic temporal expectation** *Nat Neurosci* **14**:246–51 <https://doi.org/10.1038/nn.2688> |

[PubMed](#) | [Google Scholar](#)

28. Kato HK, Gillet SN, Isaacson JS (2015) **Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance** *Neuron* **88**:1027–39 <https://doi.org/10.1016/j.neuron.2015.10.024> | [PubMed](#) | [Google Scholar](#)
29. Rothschild G, Nelken I, Mizrahi A (2010) **Functional organization and population dynamics in the mouse primary auditory cortex** *Nat Neurosci* **13**:353–60 <https://doi.org/10.1038/nn.2484> | [PubMed](#) | [Google Scholar](#)
30. Cohen L, Mizrahi A (2015) **Plasticity during motherhood: changes in excitatory and inhibitory layer 2/3 neurons in auditory cortex** *J Neurosci* **35**:1806–15 <https://doi.org/10.1523/JNEUROSCI.1786-14.2015> | [PubMed](#) | [Google Scholar](#)
31. Cohen L, Rothschild G, Mizrahi A (2011) **Multisensory integration of natural odors and sounds in the auditory cortex** *Neuron* **72**:357–69 <https://doi.org/10.1016/j.neuron.2011.08.019> | [PubMed](#) | [Google Scholar](#)
32. Maor I, Shalev A, Mizrahi A (2016) **Distinct Spatiotemporal Response Properties of Excitatory Versus Inhibitory Neurons in the Mouse Auditory Cortex** *Cereb Cortex* **26**:4242–52 <https://doi.org/10.1093/cercor/bhw266> | [PubMed](#) | [Google Scholar](#)
33. Sawyer SM, Afifi RA, Bearinger LH, Blakemore SJ, Dick B, Ezech AC, et al. (2012) **Adolescent Health 1 Adolescence: a foundation for future health** *Lancet* **379**:1630–40 [https://doi.org/10.1016/S0140-6736\(12\)60072-5](https://doi.org/10.1016/S0140-6736(12)60072-5) | [Google Scholar](#)
34. Spear LP (2000) **Neurobehavioral changes in adolescence** *Current Directions in Psychological Science* **9**:111–4 <https://doi.org/10.1111/1467-8721.00072> | [Google Scholar](#)
35. Dahl RE, Allen NB, Wilbrecht L, Suleiman AB (2018) **Importance of investing in adolescence from a developmental science perspective** *Nature* **554**:441–50 <https://doi.org/10.1038/nature25770> | [PubMed](#) | [Google Scholar](#)
36. Reynolds LM, Yetnikoff L, Pokinko M, Wodzinski M, Epelbaum JG, Lambert LC, et al. (2019) **Early Adolescence is a Critical Period for the Maturation of Inhibitory Behavior** *Cereb Cortex* **29**:3676–86 <https://doi.org/10.1093/cercor/bhy247> | [PubMed](#) | [Google Scholar](#)
37. Meyer HC, Bucci DJ (2016) **Age differences in appetitive Pavlovian conditioning and extinction in rats** *Physiol Behav* **167**:354–62 <https://doi.org/10.1016/j.physbeh.2016.10.004> | [PubMed](#) | [Google Scholar](#)
38. Magis-Weinberg L, Custers R, Dumontheil I (2019) **Rewards Enhance Proactive and Reactive Control in Adolescence and Adulthood** *Soc Cogn Affect Neurosci* **14**:1219–32 <https://doi.org/10.1093/scan/nsz093> | [PubMed](#) | [Google Scholar](#)
39. Cohen JR, Asarnow RF, Sabb FW, Bilder RM, Bookheimer SY, Knowlton BJ, et al. (2010) **A unique adolescent response to reward prediction errors** *Nat Neurosci* **13**:669–71 <https://doi.org/10.1038/nn.2558> | [PubMed](#) | [Google Scholar](#)
40. Crone EA, Dahl RE (2012) **Understanding adolescence as a period of social-affective engagement and goal flexibility** *Nat Rev Neurosci* **13**:636–50 <https://doi.org/10.1038/nrn3313> | [PubMed](#) | [Google Scholar](#)

41. Kunkhyen T, Perez E, Bass M, Coyne A, Baum MJ, Cherry JA (2018) **Gonadal hormones, but not sex, affect the acquisition and maintenance of a Go/No-Go odor discrimination task in mice** *Horm Behav* **100**:12–9 <https://doi.org/10.1016/j.yhbeh.2018.02.009> | PubMed | Google Scholar
42. Buran BN, Sarro EC, Manno FA, Kang R, Caras ML, Sanes DH (2014) **A sensitive period for the impact of hearing loss on auditory perception** *J Neurosci* **34**:2276–84 <https://doi.org/10.1523/JNEUROSCI.0647-13.2014> | PubMed | Google Scholar
43. Lee T-Y, Weissenberger Y, King AJ, Dahmen JC (2024) **Midbrain encodes sound detection behavior without auditory cortex** *eLife* **12**:RP89950 [Google Scholar](#)
44. Kato HK, Asinof SK, Isaacson JS (2017) **Network-level control of frequency tuning in auditory cortex** *Neuron* **95**:412–23 [Google Scholar](#)
45. Talwar SK, Musial PG, Gerstein GL (2001) **Role of mammalian auditory cortex in the perception of elementary sound properties** *Journal of Neurophysiology* **85**:2350–8 <https://doi.org/10.1152/jn.2001.85.6.2350> | Google Scholar
46. Haimson B, Gilday OD, Lavi-Rudel A, Sagi H, Lottem E, Mizrahi A (2024) **Single neuron responses to perceptual difficulty in the mouse auditory cortex** *Sci Adv* **10**:eadp9816 <https://doi.org/10.1126/sciadv.adp9816> | PubMed | Google Scholar
47. Ahissar M, Hochstein S (2004) **The reverse hierarchy theory of visual perceptual learning** *Trends Cogn Sci* **8**:457–64 <https://doi.org/10.1016/j.tics.2004.08.011> | PubMed | Google Scholar
48. Ahissar M, Hochstein S (1997) **Task difficulty and the specificity of perceptual learning** *Nature* **387**:401–6 <https://doi.org/10.1038/387401a0> | PubMed | Google Scholar
49. Zagha E (2020) **Shaping the cortical landscape: functions and mechanisms of top-down cortical feedback pathways** *Frontiers in Systems Neuroscience* **14** [Google Scholar](#)
50. Tasaka GI, Feigin L, Maor I, Groysman M, DeNardo LA, Schiavo JK, et al. (2020) **The Temporal Association Cortex Plays a Key Role in Auditory-Driven Maternal Plasticity** *Neuron* **107**:566–79 <https://doi.org/10.1016/j.neuron.2020.05.004> | PubMed | Google Scholar
51. Tasaka GI, Maggi C, Taha E, Mizrahi A (2023) **The local and long-range input landscape of inhibitory neurons in mouse auditory cortex** *J Comp Neurol* **531**:502–14 <https://doi.org/10.1002/cne.25437> | PubMed | Google Scholar
52. Delevich K, Klinger M, Okada NJ, Wilbrecht L (2021) **Coming of age in the frontal cortex: The role of puberty in cortical maturation** *Semin Cell Dev Biol* **118**:64–72 <https://doi.org/10.1016/j.semcdb.2021.04.021> | PubMed | Google Scholar
53. Konstantoudaki X, Chalkiadaki K, Vasileiou E, Kalemaki K, Karagogeos D, Sidiropoulou K (2018) **Prefrontal cortical-specific differences in behavior and synaptic plasticity between adolescent and adult mice** *J Neurophysiol* **119**:822–33 <https://doi.org/10.1152/jn.00189.2017> | PubMed | Google Scholar
54. Johnson CM, Loucks FA, Peckler H, Thomas AW, Janak PH, Wilbrecht L (2016) **Long-range orbitofrontal and amygdala axons show divergent patterns of maturation in the frontal**

cortex across adolescence *Dev Cogn Neurosci* **18**:113–20 <https://doi.org/10.1016/j.dcn.2016.01.005> | PubMed | Google Scholar

55. Nabel EM, Garkun Y, Koike H, Sadahiro M, Liang A, Norman KJ, et al. (2020) **Adolescent frontal top-down neurons receive heightened local drive to establish adult attentional behavior in mice** *Nat Commun* **11**:3983 <https://doi.org/10.1038/s41467-020-17787-0> | PubMed | Google Scholar
56. Druckmann S, Chklovskii DB (2012) **Neuronal circuits underlying persistent representations despite time varying activity** *Curr Biol* **22**:2095–103 <https://doi.org/10.1016/j.cub.2012.08.058> | PubMed | Google Scholar
57. Kang B, Druckmann S (2020) **Approaches to inferring multi-regional interactions from simultaneous population recordings: Inferring multi-regional interactions from simultaneous population recordings** *Curr Opin Neurobiol* **65**:108–19 <https://doi.org/10.1016/j.conb.2020.10.004> | PubMed | Google Scholar
58. Li N, Daie K, Svoboda K, Druckmann S (2016) **Robust neuronal dynamics in premotor cortex during motor planning** *Nature* **532**:459 [Google Scholar](https://doi.org/10.1038/nature16170)
59. Wei Z, Inagaki H, Li N, Svoboda K, Druckmann S (2019) **An orderly single-trial organization of population dynamics in premotor cortex predicts behavioral variability** *Nat Commun* **10**:216 <https://doi.org/10.1038/s41467-018-08141-6> | PubMed | Google Scholar
60. Wei Z, Lin BJ, Chen TW, Daie K, Svoboda K, Druckmann S (2020) **A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology** *PLoS Comput Biol* **16**:e1008198 <https://doi.org/10.1371/journal.pcbi.1008198> | PubMed | Google Scholar
61. Drieu C, Zhu Z, Wang Z, Fuller K, Wang A, Elnozahy S, et al. (2025) **Rapid emergence of latent knowledge in the sensory cortex drives learning** *Nature* :1–11 [Google Scholar](https://doi.org/10.1038/s41586-025-05000-0)
62. Takesian AE, Kotak VC, Sanes DH (2012) **Age-dependent effect of hearing loss on cortical inhibitory synapse function** *J Neurophysiol* **107**:937–47 <https://doi.org/10.1152/jn.00515.2011> | PubMed | Google Scholar
63. Takesian AE, Kotak VC, Sanes DH (2010) **Presynaptic GABA(B) receptors regulate experience-dependent development of inhibitory short-term plasticity** *J Neurosci* **30**:2716–27 <https://doi.org/10.1523/JNEUROSCI.3903-09.2010> | PubMed | Google Scholar
64. Gilad A, Maor I, Mizrahi A (2020) **Learning-related population dynamics in the auditory thalamus** *eLife* **9**:e56307 <https://doi.org/10.7554/eLife.56307> | PubMed | Google Scholar
65. Gilday OD, Praegel B, Maor I, Cohen T, Nelken I, Mizrahi A (2023) **Surround suppression in mouse auditory cortex underlies auditory edge detection** *PLoS Comput Biol* **19**:e1010861 <https://doi.org/10.1371/journal.pcbi.1010861> | PubMed | Google Scholar
66. Shani-Narkiss H, Vinograd A, Landau ID, Tasaka G, Yayon N, Terletsky S, et al. (2020) **Young adult-born neurons improve odor coding by mitral cells** *Nat Commun* **11**:5867 <https://doi.org/10.1038/s41467-020-19472-8> | PubMed | Google Scholar

67. Gilday OD, Mizrahi A (2023) **Learning-Induced Odor Modulation of Neuronal Activity in Auditory Cortex** *J Neurosci* **43**:1375–86 <https://doi.org/10.1523/JNEUROSCI.1398-22.2022> | PubMed | Google Scholar
68. Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, et al. (2017) **Fully integrated silicon probes for high-density recording of neural activity** *Nature* **551**:232–6 <https://doi.org/10.1038/nature24636> | PubMed | Google Scholar
69. Pachitariu M, Steinmetz N, Kadir S, Carandini M, Kenneth D H. (2016) **Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels** *bioRxiv* :061481 <https://doi.org/10.1101/061481> | Google Scholar
70. Shamash P, Carandini M, Harris K, Steinmetz N (2018) **A tool for analyzing electrode tracks from slice histology** *BioRxiv* <https://doi.org/10.1101/447995> | Google Scholar
71. Vinje WE, Gallant JL (2000) **Sparse coding and decorrelation in primary visual cortex during natural vision** *Science* **287**:1273–6 <https://doi.org/10.1126/science.287.5456.1273> | PubMed | Google Scholar
72. Green DM (1966) **Swets JA** *Signal detection theory and psychophysics: Wiley New York* | Google Scholar

Author information

Benedikt Praegel

The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel, Department of Neurobiology, the Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
ORCID iD: [0009-0006-8253-9625](https://orcid.org/0009-0006-8253-9625)

Feng Chen

Department of Neurobiology, Stanford University, Stanford, United States, Department of Applied Physics, Stanford University, Stanford, United States

Adria Dym

The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

Amichai Lavi-Rudel

The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

Shaul Druckmann

Department of Neurobiology, Stanford University, Stanford, United States, Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, United States
ORCID iD: [0000-0003-0068-3377](https://orcid.org/0000-0003-0068-3377)

Adi Mizrahi

The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel, Department of Neurobiology, the Institute of Life Sciences, The Hebrew

University of Jerusalem, Jerusalem, Israel

ORCID iD: [0000-0002-1743-6754](https://orcid.org/0000-0002-1743-6754)

For correspondence: Mizrahi.adi@mail.huji.ac.il

Editors

Reviewing Editor

Brice Bathellier

Centre National pour la Recherche Scientifique et Technique (CNRS), Paris, France

Senior Editor

Michael Frank

Brown University, Providence, United States of America

Reviewer #1 (Public review):

Summary:

Praegel et al. explore the differences in learning an auditory discrimination task between adolescent and adult mice. Using freely-moving (Educage) and head-fixed paradigms, they compare behavioral performance and neuronal responses over the course of learning. The mice were initially trained for seven days on an easy pure frequency tone Go/No-go task (frequency difference of one octave), followed by seven days of a harder version (frequency difference of 0.25 octave). While adolescents and adults showed similar performance on the easy task, adults performed significantly better on the harder task. Quantifying the lick bias of both groups, the authors then argue that the difference in performance is not due to a difference in perception, but rather to a difference in cognitive control. The authors then used neuropixel recordings across 4 auditory cortical regions to quantify the neuronal activity related to the behavior. At the single cell level, the data shows earlier stimulus-related discrimination for adults compared to adolescents in both the easy and hard tasks. At the neuronal population level, adults displayed a higher decoding accuracy and lower onset latency in the hard task as compared to adolescents. Such differences were not only due to learning, but also to age as concluded from recordings in novice mice. After learning, neuronal tuning properties had changed in adults but not in adolescent. Overall, the differences between adolescent and adult neuronal data correlates with the behavior results in showing that learning a difficult task is more challenging for younger mice.

Strengths:

The behavioral task is well designed, with the comparison of easy and difficult tasks allowing for a refined conclusion regarding learning across age. The experiments with optogenetics and novice mice are completing the research question in a convincing way.

The analysis, including the systematic comparison of task performance across the two age groups, is most interesting and reveals differences in learning (or learning strategies?) that are compelling.

Neuronal recording during both behavioral training and passive sound exposure is particularly powerful, and allows interesting conclusions.

Weaknesses:

The weaknesses listed by this reviewer were addressed by adequate revisions.

<https://doi.org/10.7554/eLife.106387.3.sa2>

Reviewer #2 (Public review):

Summary:

The authors aimed to find out how and how well adult and adolescent mice discriminate tones of different frequencies and whether there are differences in processing at the level of the auditory cortex that might explain differences in behavior between the two groups. Adolescent mice were found to be worse at sound frequency discrimination than adult mice. The performance difference between the groups was most pronounced when the sounds are close in frequency and thus difficult to distinguish and could, at least in part, be attributed to the younger mice' inability to withhold licking in no-go trials. By recording the activity of individual neurons in the auditory cortex when mice performed the task or were passively listening as well as in untrained mice the authors identified differences in the way that the adult and adolescent brains encode sounds and the animals' choice that could potentially contribute to the differences in behavior.

Strengths:

The study combines behavioural testing in freely-moving and head-fixed mice, optogenetic manipulation and high density electrophysiological recordings in behaving mice to address important open questions about age differences in sound-guided behavior and sound representation in the auditory cortex.

Weaknesses:

The weaknesses listed by this reviewer were addressed by adequate revisions.

<https://doi.org/10.7554/eLife.106387.3.sa1>

Author response:

The following is the authors' response to the previous reviews.

Reviewer #1 (Public review):

Summary:

Praegel et al. explore the differences in learning an auditory discrimination task between adolescent and adult mice. Using freely-moving (Educage) and head-fixed paradigms, they compare behavioral performance and neuronal responses over the course of learning. The mice were initially trained for seven days on an easy pure frequency tone Go/No-go task (frequency difference of one octave), followed by seven days of a harder version (frequency difference of 0.25 octave). While adolescents and adults showed similar performance on the easy task, adults performed significantly better on the harder task. Quantifying the lick bias of both groups, the authors then argue that the difference in performance is not due to a difference in perception, but rather to a difference in cognitive control. The authors then used neuropixel recordings across 4 auditory cortical regions to quantify the neuronal activity related to the behavior. At the single cell level, the data shows earlier stimulus-related discrimination for adults compared to adolescents in both the easy and hard tasks. At the neuronal population level, adults displayed a higher decoding accuracy and lower onset latency in the hard task as compared to adolescents. Such differences were not only due to learning, but also to age as concluded from recordings in novice mice. After learning, neuronal tuning properties had changed in adults but not in adolescent. Overall, the differences between adolescent

and adult neuronal data correlates with the behavior results in showing that learning a difficult task is more challenging for younger mice.

Strengths:

The behavioral task is well designed, with the comparison of easy and difficult tasks allowing for a refined conclusion regarding learning across age. The experiments with optogenetics and novice mice are completing the research question in a convincing way.

The analysis, including the systematic comparison of task performance across the two age groups, is most interesting, and reveals differences in learning (or learning strategies?) that are compelling.

Neuronal recording during both behavioral training and passive sound exposure is particularly powerful, and allows interesting conclusions.

Weaknesses:

The presentation of the paper must be strengthened. Inconsistencies, missing information or confusing descriptions should be fixed.

We have carefully re-read the manuscript and reviewed it for inconsistencies. We made several corrections in the figures. For example, we removed redundant lines from violin plots and statistics, applied consistent labels, matched y- and x-limits of graphics, and adjusted labels. We also clarified descriptions of some experiment by adding explanations to the text.

The recording electrodes cover regions in the primary and secondary cortices. It is well known that these two regions process sounds quite differently (for example, one has tonotopy, the other not), and separating recordings from both regions is important to conclude anything about sound representations. The authors show that the conclusions are the same across regions for Figure 4, but is it also the case for the subsequent analysis? Comparing to the original manuscript, the authors have now done the analysis for AuDp and AUDv separately, and say that the differences are similar in both regions. The data however shows that this is not the case (Fig S7). And even if it were the case, how would it be compatible with the published literature?

To address this and previous concerns about regional differences, the manuscript now includes 4 figures (4-1, 4-3, 6-2, 7-1) and 5 supplemental tables (3,4, 5, 6, 8) that explicitly compare results across brain regions.

Following the reviewer's request for subsequent analysis, we now added a new supplemental figure (Fig. S6-2) and two new supplementary tables (Tables S5, S6). We show that similar to expert mice (supplementary Table 3, and supplementary Table 4), the firing properties of adolescent and adult novice mice differ across auditory subregions (supplementary Table 5). We also show that the different auditory subregions have different firing properties (supplementary Table 6). With respect to task engagement, we show that (similar to Fig. S4-2) the neuronal discriminability in different auditory subregions is similar in both novice and expert mice (Fig. S6-2).

Following the comment on Fig. S7-1, we made three changes to the revised manuscript. First, we now highlight that the differences firing properties between adolescent and adult neurons in AUDp and AUDv were distinct, but not significantly different within age-group comparisons. Second, we clearly state that the learning related changes in the measured parameters are different between AUDp and AUDv. Note, however, the greater changes in adult neurons after learning remains consistent between AUDp and AUDv. Third, we softened our original claim but still highlighted the stronger learning-induced plasticity in adults.

Regarding the concern that different regions should show different patterns due to their known differences (e.g. tonotopy). Of course we agree that different areas differ functionally (as shown in our own previous work and here as well). However, it is still plausible, and biologically reasonable, that developmental changes may proceed in a similar direction across different areas, even if their baseline coding properties differ.

Reviewer #2 (Public review):

Summary:

The authors aimed to find out how and how well adult and adolescent mice discriminate tones of different frequencies and whether there are differences in processing at the level of the auditory cortex that might explain differences in behavior between the two groups. Adolescent mice were found to be worse at sound frequency discrimination than adult mice. The performance difference between the groups was most pronounced when the sounds are close in frequency and thus difficult to distinguish and could, at least in part, be attributed to the younger mice' inability to withhold licking in no-go trials. By recording the activity of individual neurons in the auditory cortex when mice performed the task or were passively listening as well as in untrained mice the authors identified differences in the way that the adult and adolescent brains encode sounds and the animals' choice that could potentially contribute to the differences in behavior.

Strengths:

The study combines behavioural testing in freely-moving and head-fixed mice, optogenetic manipulation and high density electrophysiological recordings in behaving mice to address important open questions about age differences in sound-guided behavior and sound representation in the auditory cortex.

Weaknesses:

For some of the analyses that the authors conducted it is unclear what the rationale behind them is and, consequently, what conclusion we can draw from them.

We have carefully re-read the manuscript and reviewed it for analyses that lacked a clear rationale or conclusion. To address this, we have made several changes to clarify the reasoning and strengthen the interpretation of the results.

Reviewer #1 (Recommendations for the authors):

It would have helped if the authors had highlighted the changes they made to the manuscript compared to the original version - especially since many replies to the reviewers' comments were as vague as "...we fixed some of the wording so it adheres to the data shown", or "we refined our interpretation", without further details.

The revised version has improved substantially, and the main claims have been discussed in a more objective way. Important new analyses have been added to allow for a refined interpretation of the results. However, the presentation of the data could still be strengthened significantly (in response to comment A from last review).

We apologize for the lack of detail in some of our previous responses. Our intention was to keep the replies concise, assuming that the side-by-side version with tracked changes would make the edits sufficiently clear. However, we understand the need for greater transparency. Thus, below we provide the following five lists describing the major changes: (1) List of specific reviewer recommendations, (2) list of corrections in figures, (3) list of clarity issues,

(4) list of fixed mistakes, (5) list of new figures. We hope this breakdown makes the revisions clearer and more accessible.

List of specific reviewer recommendations:

I.108 mentions a significant change in the vertical line of Fig 1F - Could this significance be indicated and quantified in the figure?

We quantified and indicated the significance of the vertical line in Fig. 1f and Fig. 1i.

Fig.1G - the thick and thin lines should be defined, as well as the grey and white dots (same values for adolescents, not for adults).

(a) We removed the thin inner lines from the violin plot. We define the bar (thick line) of the violin plot in an additional sentence in the methods section under data analysis (LL820-823).
 (b) We adjusted the marker outlines in the adult data (Fig. 1G).

the figure axis legends should be consistent (trials in Fig D vs # trials in Fig 1F)

We adjusted the axis legend to # trials in Fig. 1D.

I.110: is d' always calculated based on the 100 last trials of a session, or is it just for Figure 1F? -etc...

d' is always calculated based on the last 100 trials. To clarify this, we added a description in the methods section (L830).

List of corrections in the figures:

- (1) We removed the internal lines from violin plots in throughout Fig. 1-7.
- (2) We removed the underline of the statistics throughout Fig. 1-7.
- (3) We consistently applied 'adolescent' and 'adult' figure labels and titles with lowercase letters throughout Fig. 1-7.
- (4) We applied consistent labelling of 'time (ms)' throughout Fig. 1-7.
- (5) We matched the size of dashed lines throughout Fig. 1-6.
- (6) We adjusted the x-label of Fig. 1d, Fig. S1-1 a, Fig. 3c, Fig. 3h-i, Fig. 4d to '# trials'.
- (7) We removed the x-label of 'Experimental Group' from Fig. 1 to enhance consistency with other figures.
- (8) We removed misaligned dots from the violin plots in Fig. 1g, Fig. 2f, Fig. 3f,g.
- (9) We corrected the plot in Fig. S1-1b.
- (10) We adjusted the y-limits of Fig. S1-1c to be consistent with Fig. S1-1d,e.
- (11) We adjusted the x-labels and y-labels of Fig. 2, Fig. S3-1, Fig. S3-2 and Fig. 3b to 'freq. (kHz)'.
- (12) We added the age of adolescent and adult mice to the schematic timeline in Fig. 2a.
- (13) We added a label of the reinforcement delay to the schematic trial structure in Fig. 3b.
- (14) We added within-group statistics to Fig. 3e and the figure legend.

- (15) We adjusted the x-label of Fig. 3d to ‘# sessions’.
- (16) We adjusted the x-label of Fig. 3d and Fig. S3-1b to ‘# licks’.
- (17) We changed the y-label in Fig. S3-1a, and Fig. S3-2d, e to ‘lick ratio’ to avoid confusion with the lick rate (Hz) that was calculated in Fig. 4 and Fig. 6.
- (18) We replaced the titles ‘CAMKII’ with ‘dTomato’ in Fig. S3-2 to correctly highlight that both the experimental and control injection were CAMKII injections.
- (19) We adjusted the x-labels and y-labels of Fig. 2, Fig. S3-1, Fig. S3-2 and Fig. 3b to ‘freq. (kHz)’.
- (20) We adjusted the y-label of Fig. S4-1c to ‘# neurons’.
- (21) We matched the x-ticks in Fig. 4e,f.
- (22) We matched the x-ticks in Fig. 6d-g.
- (23) We changed the x-label in Fig. 4g, S4-2 and S6-2 to ‘duration (ms)’ to match the figure label with the manuscript.
- (24) We consistently label ‘Hit’, ‘Miss’, ‘FA’ and ‘CR’ with capital letters in Fig. 4d-e.
- (25) We replaced the double figure label ‘C.’ in Fig. S4-2 with ‘D.’.
- (26) We adjusted the dot-size in Fig. 5 to be equal for all graphs.
- (27) We added ticks to the experimental timeline in Fig. 6a.
- (28) We corrected the y-label in Fig. 7c. Now it correctly reflects 5 attenuations from 72-32 dB SPL.
- (29) We matched the y-label of Fig. 7e-h and Fig. S7-1.

List of clarity issues:

- (1) We replaced the term ‘lower response bias’ with ‘higher lick bias’ (L24) to accurately describe the more negative (lower) criterion-bias, which highlights a higher tendency to lick.
- (2) We replaced the term ‘response bias’ with ‘lick bias’ to consistently describe the calculated criterion-bias (L24, L149, L164, L455, L456, L468).
- (3) We clarify that the age-related differences were ‘more pronounced’ instead of simply ‘higher’ to accurately reflect not simply the increase in adolescent lick-bias, but also the decrease in adult lick-bias (L31).
- (4) We clarified that adolescent sound representations are not merely ‘distinct’, but ‘not fully mature’ in L83.
- (5) We clarified in L180 that the impulsive responses we observed in adolescent mice could be related to being ‘less impacted by punishments’.
- (6) We clarified the differences in firing properties of auditory sub-regions analyzed in Supplementary Table 3 (L287-295).
- (7) We explained and clarified the reference to Fig. 3j (LL252-253).
- (8) We added statistics to Fig. S4-2 to support our claim that there are no differences in the onset-latency, duration of discriminability and maximal discriminability between different

sub-regions within age-groups (LL 314-315).

(9) We expanded our explanation of the results in Table 3 (LL370-379).

(10) We separated the reference to Fig. 6b and Fig. 6c to clarify their meaning (LL358-361).

(11) We clarified the differences in basic firing properties during the FRA protocol in Fig. 7 (LL409-418).

(12) We expanded our explanation of the differences of the learning related firing properties in AUDp and AUDv of Fig. S7-1 (LL426-433).

(13) We changed the term ‘plasticity profiles’ to ‘learning related plasticity’ to further clarify our limitation that L5/6 and L2/3 may exhibit distinct learning related changes (L496).

(14) We changed the term ‘sluggish’ (L481) to ‘delayed’ to more precisely explain differences between adolescent and adult tuning properties.

(15) We clarified that the running d' was calculated in bins of 25 trials, instead of ‘the last 25 trials’ (LL845-846).

List of fixed mistakes:

(1) We corrected and matched the age to more accurately reflect the age mice were recorded (P37-42 and P77-82).

(2) We corrected the attenuation range from 72-42 to 72-32 dB SPL to correctly reflect the 5 attenuations used in the protocol.

(3) We corrected the number of channels shown in the voltage trace from 10 to 11 (Fig. S4-1a)

(4) We corrected the number of neurons recorded in novice adolescent mice in the legend of Fig. 6 from 140 to 130 (Fig. 6b).

(5) We removed redundant, or double brackets, commas, dots, and semi-colons in the figure legends.

(6) We corrected the LME statistics Table 2.

List of new figures and tables:

(1) We added a new supplementary figure to accompany Figure 6. Specifically, Fig. S6-2, shows the interaction of the three measured discriminability properties (onset delay, duration of discriminability, and maximal discriminability) in novice compared to expert mice in the easy and hard task (Go compared to No Go). The figure compares the different auditory sub-regions (similar to Fig. S4-2). We show that the discriminability properties within different groups is not significantly different among the four different sub-regions.

(2) Supplementary Table 5: We compared the firing properties in different auditory subregions in novice mice, and found (similar to expert mice) that the firing properties differ between adult and adolescent mice across the four different sub-regions.

(3) Supplementary Table 6: We compared the firing properties between different subregions, separately for adolescent and adult novice mice. Similar to expert mice, we found that different auditory subregions differ in their auditory firing properties.

Reviewer #2 (Recommendations for the authors):

The authors largely addressed my suggestions.

*Comparing hit vs correct rejection trials in the population decoding analysis (L313-314):
The authors acknowledge that comparing these two trial types conflates choice and stimulus decoding but I am not convinced that the changes to the manuscript text make this clear enough to the reader.*

Thank you for pointing this out. We have made additional revisions to clarify this, and other issues more explicitly, as follows:

(1) We have expanded the explanation of how our population decoding analysis conflates stimulus and choice, and we acknowledge the limitations of this approach in the Abstract (L28), the Results section (L324-326, LL367-370) and the Discussion (LL516-519).

(2) We replaced the analysis of impulsivity on the head-fixed task. Instead of analyzing all it is, we focus only on ITIs following FA trials (Fig. S3-1c,d). This is more consistent with the analysis in the Educage (Fig. S2-1), where we show that adolescents exhibit increased impulsivity after FA trials. We found a similar result for ITIs following FA trials in the head-fixed task.

(3) To provide complementary insight, we now further justify our use of the Fisher separation metric alongside decoding accuracy in Figure 5, with a clearer rationale provided in LL343-345

(4) We also clarified our reasoning for focusing on 62 dB SPL in the FRA-based analysis in LL400-403.

<https://doi.org/10.7554/eLife.106387.3.sa0>