Age and Learning Shapes Sound Representations in Auditory Cortex During Adolescence

Benedikt Praegel[1,2], Feng Chen[3,4], Adria Dym[1], Amichai Lavi-Rudel[1], Shaul Druckmann[3,5], and Adi Mizrahi[1,2,*]

[1]The Edmond and Lily Safra Center for Brain Sciences. [2]Department of Neurobiology, the Institute of Life Sciences. The Hebrew University of Jerusalem, Jerusalem, Israel. [3]Department of Neurobiology, Stanford University, Stanford, California, USA. [4]Department of Applied Physics, Stanford University, Stanford, California, USA. [5]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California, USA.

** Corresponding author: Adi Mizrahi: Mizrahi.adi@mail.huji.ac.il*

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.

Keywords:

Adolescence, Auditory Cortex, Auditory Learning, Neuropixels, Neuronal Coding

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

113 pure tones (**Fig. 1b**; 'Easy + Hard'). Immediately after introducing the hard pair, the
114 performance of adolescent mice on the easy task dropped, while those of adults remained
115 unaffected (**Fig. 1g**). During the second week of training, adolescents regained their high d'
116 values on the easy task, such that by the end of the experiment both groups had similarly high
117 performance on the easy task (**Fig. 1h).**

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

127 Each mouse performed a unique number of trials, as these were initiated by the mice
128 spontaneously. As noted above, pre-adolescent mice performed fewer trials/day as compared
129 to adults but, notably, only during the first week of training (mean ± STE: adolescents, 8317 ±
130 712 trials, adults 13229 ± 514 trials, t-stat = -5.6314, df = 28, p = 4.9566e-06, two-sample
131 independent t-test). Still, this difference raises the question whether the different number of
132 trials affected our conclusions from the first week of training. Thus, to evaluate the possible
133 connection between the number of trials and performance in our data, we carried out several
134 analyses. First, we found that the number of trials and d' were not correlated in either age-group
135 during the first week of discrimination on the easy task (Pearson-r; adolescent: r = 0.1880, p =
136 0.5023; adult: r = -0.0730, p = 0.7959), nor on the second week of discrimination on both the
137 easy and the hard task (Pearson-r; adolescent: r = 0. 2264, p = 0. 4171; adult: r = 0. 0194, p =
138 0. 9454). Second, we compared performance at three time points along the task with reference
139 to times with shared number of trials: 1) evaluating d' after the minimal number of trials
140 performed by all mice during the first week of training, and 2) evaluating d' at the mean number

141 of trials of each group as a reference. We found no significant differences between adolescents
142 and adults in all comparisons (**Fig. S1-1c-e**). Additionally, we found no significant differences
143 between males and females, nor among experiments that had different numbers of co-housed
144 mice in the Educage (**Table 1**). Taken together, despite the different absolute number of trials
145 during the first week, we conclude that adolescents and adults learned the easy task equally
146 quickly and effectively. The central difference between the age groups was in their ability to
147 learn the hard task during the second week, when adults outperformed adolescents.
148
149 *Adolescents have higher lick bias and higher behavioral variability*
150 We next asked what perceptual and/or cognitive aspect of behavior is different among the age
151 groups. We plotted psychometric curves based on both learned and probe trials from the last
152 2000 trials of the experiment (**Fig. 2a**). The false alarm rate but not the hit rate of adolescent
153 mice was significantly higher as compared to adults (**Fig. 2b**). To calculate the decision
154 boundary of mice and their perceptual sensitivity we normalized the psychometric curves of
155 each mouse using a unity-based normalization and sigmoid fitting (**Fig. 2c**). We found no
156 differences in the decision boundary or slopes of the curves, suggesting that perceptual
157 sensitivity is not different between the groups (**Fig. 2c**). To test whether the differences arise
158 from a cognitive effect (e.g., lick bias), we calculated the decision threshold as the maximal
159 lick bias, also known as the criterion bias (C-bias). C-bias reflects the tendency to respond in a
160 liberal (i.e., negative C-bias) or a conservative (i.e., positive C-bias) manner to the sounds
161 during the task. C-bias was negative for both groups, yet significantly lower in adolescents,
162 suggesting that their cognitive ability to withhold licking is inferior to that of adults (**Fig. 2d**).
163 Given that adolescents and adults show different dynamics in their performance along
164 the task (**Fig. 1**), we tested whether the differences we found in the lick bias are consistent at
165 different times along the task. To normalize for different numbers of trials in each mouse, we
166 divided their training episodes into the first and last tertiles during the first week and second
167 week of training separately (**Fig. 2e**). The maximal lick bias was significantly higher in
168 adolescents in all but the very first training episode, when mice learned the procedure (**Fig. 2f**).

169 These data highlight the general cognitive difficulty of adolescents to withhold licking during
170 the task.

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

182 Prior studies have shown that a dominant feature of adolescent behavior is that it is
183 more variable compared to adults[4]. To examine behavioral variability in our data, we
184 calculated the coefficient of variation (CV) of the discriminability (d') for each mouse across
185 training. While there was no significant difference in d' variability when training on the easy
186 task during the first week of training (**Fig. 2h**), adult mice showed significantly lower variability
187 during the second week of training while training on the 'easy + hard' versions of the task (**Fig.**
188 **2i,j**). Thus, adolescent learning and behavior are characterized by both lower response
189 inhibition and higher variability, particularly during the second week when the task involved
190 more challenging discrimination for the younger mice.

191

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

193 To enable electrophysiological recordings during expert task performance, we trained
194 adolescent (n = 5) and adult (n = 6) mice on a head-fixed learning paradigm, using a similar
195 protocol as in the Educage (**Fig. 3a**). In the head-fixed protocol, water access was limited to
196 training sessions. Head-fixed mice were trained to lick after a 100ms tone and were rewarded

197 (with water) or punished (by a 2 sec white noise) only after a 600ms delay (**Fig. 3b**). Each
198 session was concluded after the mice became satiated and stopped licking. Water supply was
199 limited to 0.0125 ml per day per gram of body weight, and mice that did not consume this
200 amount were compensated after training (see Methods; **Fig. 3c** shows an example session from
201 one mouse). The number of trials per training session was significantly lower in adolescents
202 (mean ± STE: adolescents: 410 ± 35, adults: 580 ± 37 t-statistic = -27.9855, p = 0.0012,
203 independent two-sample t-test). However, the number of sessions to reach the behavioral
204 threshold (d' >= 1 in the easy task) was not significantly different between adolescents and
205 adults (mean ± STE: adolescents: 3.7 ± 0.8, adults: 3.8 ± 0.7; t-statistic = - 0.8531, p = 0.1971,
206 independent two-sample t-test, **Fig. 3d** shows all learning curves). Similar to learning in the
207 Educage, adolescent performance (but not that of adults) decreased after we introduced the hard
208 task (**Fig. 3e**). Another similarity between the head-fixed and Educage versions of the task was
209 that we found no significant differences between adolescents and adults in the performance of
210 the easy task (**Fig. 3f**), and that adults outperformed adolescents on the hard discrimination
211 (**Fig. 3g**). Also similar to the behavior in the Educage, lick responses of adolescent mice were
212 generally higher (**Fig. S3-1a**), and adolescents exhibited a stronger lick bias (**Fig. S3-1b**). In
213 line with this result, in the head-fixed task we found higher impulsiveness in the proportion and
214 the number of licks during the inter-trial intervals after FA trials in adolescents (**Fig. S3-1c, d**).
215 To test whether ACx was necessary for expert discrimination in this task, we performed
216 a causal experiment in adult mice. Adult mice (n=3) were injected bilaterally with AAV5-
217 CAMKII-GtACR2-FRED_kv_2.1 into the ACx, and optical fibers were implanted over the
218 injection sites (**Fig. S3-2a, c**; **Fig. S3-3**). This viral construct enables optogenetic silencing of
219 excitatory neurons via GtACR2 (*Guillardia theta* Anion Channelrhodopsin 2), under the
220 control of the CaMKII promoter. It includes a red fluorescent protein for expression verification
221 (FRED) and a soma-targeting motif derived from the Kv2.1 potassium channel to enhance
222 membrane localization.

223 Mice were trained identically on the head-fixed Go/No-Go paradigm as described
224 above. At the end of training, we attached a light source to the implanted fibers to allow
225 optogenetic suppression during several sessions (**Fig. S3-2a**). In these sessions, light was
226 applied in 50% of trials in a pseudo-random fashion, with light-on trials starting from -50ms
227 prior to tone onset up to 50ms after tone offset (**Fig. S3-2b**). A control group of mice (n= 3)
228 were injected with AAV9-CaMKII-dTomato and went through the exact same procedure (**Fig.**
229 **S3-2a, c**). We compared the lick responses for both Go and No-Go stimuli under light-on and
230 light-off conditions. We found that lick responses were strongly affected in the GtACR2-
231 injected mice (n = 13 sessions), but not in controls (n = 8 sessions) (**Fig. S3-2d**). Light-on
232 conditions only affected the lick responses of the inhibited trial and not the light-off trials (**Fig.**
233 **S3-2e**). Discriminability (d') decreased significantly with optogenetic suppression for both easy
234 and hard tasks, although all mice still performed better in the easy task under both conditions
235 (**Fig. S3-2f**). These results confirm findings by others in similar (though not identical) tasks
236 [19, 20], suggesting that the ACx is necessary for executing the behavior during the expert stage
237 of this task (but see Discussion for limitations of this experiment).

238 To measure neural responses during behavioral performance, head-fixed expert mice
239 underwent multiple recording sessions targeting the ACx (adults, n = 14 recording sessions
240 from 6 mice, age: P77-P82; adolescents, n = 13 recordings from 5 mice, age: P37-P42). During
241 the recordings, the performance of all mice on the easy task was at least d'>1, and the d' on the
242 hard task was smaller for both groups (often smaller than 1; **Fig. S3-4a**). Performance was not
243 significantly different across the multiple recording sessions in each mouse (**Fig. S3-4b**).
244 Behavioral performance during recordings was not significantly different to performance in the
245 Educage (**Fig. S3-4c**). **Figure 3h** shows the evolution of behavioral discriminability (d') during
246 all recordings for adolescents and adults. The performance was heterogeneous across mice as
247 well as within recording-sessions. However, the total number of trials performed was not
248 significantly different between age groups (trials per recording: adolescent: 340.5 ± 45; adult:
249 431.1 ± 30.3; independent t-test, t-statistic = -203.7581, p = 0.1116).

250 We found additional differences between the behavior of adults and adolescents in the
251 head restrained configuration. The average adolescent discriminability was lower at the
252 beginning of the session (initial 78 trials) but then rapidly leveled out (**Fig. 3i**). **Figure 3j** shows
253 the cumulative lick curves of adults and adolescents. Adolescent mice had shorter lick latencies
254 (**Fig. 3k**), and higher lick counts (**Fig. 3l; Supplementary Table 1**). While the lick latencies
255 were independent of the discriminability (d'), we found a significant interaction effect between
256 lick count and the d' (**Supplementary Table 1**). Together, these behavioral differences suggest
257 that the weaker response inhibition (here, expressed as lick latency) and higher reward
258 anticipation (here, expressed as lick count) may contribute to the weaker performance of
259 adolescents during the beginning of the recording session.
260
261 *Neuronal representations of stimulus- and choice-related activity are immature in adolescents*
262 To compare the activity of neurons in adolescent and adult mice, we recorded spiking activity
263 from the ACx using the high-density Neuropixels-1 probes in expert mice engaged in the task.
264 We targeted the ACx by inserting the probe in a diagonal angle, traversing four auditory
265 regions: Dorsal Auditory Cortex (AUDd), Primary Auditory Cortex (AUDp), ventral auditory
266 cortex (AUDv), and Temporal Association Cortex (Tea; **Fig. 4a; Fig. S4-1a,b,** show an
267 example voltage trace and one example recording). We recorded multiple times from each
268 mouse and used DiI- or DiO-coatings for different penetrations to verify probe trajectories
269 postmortem. We used the 3D-allen CCF-slice reconstruction platform (github.com/cortex-
270 lab/allenCCF), for high resolution anatomical registration (**Fig.4b**; **Fig. S4-1c, d**).
271 We limited our single neuron analysis to well-isolated single units from infragranular
272 layer 5 and layer 6 (recordings were largely restricted to these layers due to the probe angle),
273 as two main projection layers of the cortex, which are key nodes for assessing cortical outputs
274 [21]. Recorded neurons from other layers and areas were excluded from the analysis. Together,
275 we collected data from 1145 single neurons in 13 recordings of 5 adolescent mice and 1267
276 single neurons in 14 recordings of 6 adult mice (**Supplementary Table 2;** an overview of every
277 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 -200ms to 600ms 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.

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

306 AUC values over time from an example neuron (same neuron shown in **Fig. 4D**), and **figure 4f**
307 shows the average calculated from all neurons. The vast majority of neurons successfully
308 discriminated (AUC above shuffled data) either or both stimulus- and choice-related activity
309 (adolescent: easy stimuli = 93%, hard stimuli = 93%, choice= 95%; adult: easy stimuli = 97%,
310 hard stimuli = 97%, choice= 95%). The onset-latency of discriminability was significantly
311 slower, maximal discriminability significantly weaker, and the duration of discriminability
312 significantly shorter in adolescent neurons compared to adult neurons (discriminability traces
313 are plotted in **Fig. 4f**, and all values from individual neurons and the 3 metrics are plotted
314 together in **Fig. 4g**; see **Table 2** for statistics). The four auditory subregions were not
315 significantly different in all 3 measured metrics (**Fig. S4-2**). Thus, auditory processing of
316 stimulus- and choice-related activity in the adolescent ACx has not yet reached full maturity.
317
318 *Adults exhibit better population decoding of the difficult task compared to adolescents*
319 While individual neurons exhibited weaker discriminability in adolescents than in adults, at a
320 population level these effects could saturate out. We therefore tested the ability of recorded
321 populations to decode trial outcomes. We used a linear discriminant analysis (LDA) decoder to
322 quantify hit versus correct reject trial outcomes. We applied stringent inclusion criteria for the
323 number of simultaneously recorded neurons (see Methods), which restricted our analysis to hit
324 versus correct rejection trials with sufficient trial counts for statistical reliability. We note that
325 this comparison extracts information of the combination of stimulus and choice information
326 indiscriminately (i.e. collectively it includes the full 'task-related' differences). We decoded
327 separately in easy trials and hard trials, based on population activity during the first 200
328 milliseconds. For the easy task there was no significant difference between decoding accuracy
329 in adolescents and adults (**Fig. 5a**). However, decoding accuracy was significantly higher in
330 adults compared to adolescents in hard trials (**Fig. 5a**). For both adults and adolescents,
331 decoding accuracy was significantly lower in hard trials compared to easy trials (**Fig. 5a**). We
332 performed finer timescale decoding to quantify the onset latency of differences between the
333 activity for hit vs. correct reject trials, defined as the time in which the decoding accuracy first

334 crosses three times the standard deviation away from the baseline (see Methods). We found that
335 the latency was shorter in adults in both the easy and hard trials (**Fig. 5b**). In adolescent animals
336 the latency was larger in hard trials (**Fig. 5b**) but not in adult animals (**Fig. 5b**). We then
337 extended this analysis over a temporal window from -0.5s to10s relative to tone onset (**Fig. 5c**).
338 In the 200ms window after the response period (2.0-2.2 s), there is no longer a significant
339 difference in decoding accuracy between adult and adolescent mice as well as between the two
340 task difficulties (**Fig. S5-1**).

341 To further quantify the different decoding performance between two tasks difficulties,
342 we plotted the fisher separation metric ($\frac{\sigma^2_{between}}{\sigma^w_{within}}$, see Methods) for each level of difficulty. While
343 decoding accuracy reflects how well a classifier can distinguish between trial types, the Fisher
344 separation metric specifically quantifies the ratio of between-category variance to within-
345 category variance. The separation ratio of hard over easy tasks was significantly higher in adult
346 mice than in adolescent mice (robust linear regression, see Methods). Separation could have
347 changed due to alterations in the mean or the dispersion around the mean. Testing the
348 dispersion, we found that the variance around the means was not equal across easy and hard
349 trials but rather significantly increased for the hard trials (**Fig. 5d-f**). However, there was no
350 significant difference in variance between adolescent and adult mice within the same task **(Fig.**
351 **5d-f**).

352

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

354 We next studied how learning contributes to cortical plasticity in the different age groups. To
355 separate between age- and learning-related effects, we recorded from the ACx of two new
356 groups of mice (adolescents and adults that are age-matched to expert mice) that are novice on
357 the task (**Fig. 6a**, novice P37-P42, recordings= 6; novice P77-P82, recordings =6, n = 3 mice
358 per group). We collected data from 130 tone-responsive (by excitation) neurons in novice
359 adolescents and 186 tone-responsive (by excitation) neurons in novice adults (**Fig. 6b**;
360 **Supplementary Table 2**). **Figure 6c** shows the normalized average FR PSTH together with

361 the LR PSTH for novice mice. Notably, novice mice did not discriminate between sounds (**Fig.**
362 **S6-1a, b**). Similar to expert mice **(Supplementary Table 3**), we found diverse firing properties
363 across auditory subregions in novice mice (**Supplementary Table 5**). Also similar to expert
364 mice, several basic firing properties were different across regions in both novice adolescents
365 and novice adults (**Supplementary Table 6**).

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

380 To compare neuronal discriminability and behavior more explicitly, we analyzed the
381 correlations between neuronal decoding and behavioral performance using linear regression. In
382 the easy task, both age groups showed high correlation between neuronal discriminability and
383 behavioral discriminability (**Fig. 6h,i**; light blue). In the hard task, only adult neurons had
384 significant correlations between neuronal discriminability and behavioral discriminability (**Fig.**
385 **6h,i**; dark blue).

386

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

388 Auditory learning has been shown to induce changes in the tuning properties of sounds in the
389 ACx [18, 19, 21, 23-28]. To study whether learning-induced plasticity in ACx is distinct in
390 adolescents and adults, we recorded from the same mice described above under passive
391 listening conditions by simply extending the recording following the engaged session (**Fig. 7a,**
392 **Supplementary Table 7**; Adolescents recordings = 4; Adults recordings = 4). These data were
393 compared to data from a group of novice mice, recorded under passive listening conditions
394 following their own engaged session (**Supplementary Table 7**; novice adolescent recordings
395 =6; novice adult recordings =6).

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

404 Our dataset included four groups. The two adolescent groups were: 1) adolescent
405 novices, and 2) adolescent experts, with 92 and 80 neurons, respectively. The two adult groups
406 were: 3) adult novices, and 4) adult experts, with 123 and 84 neurons, respectively. Neurons in
407 all groups responded to pure tone as expected from previous studies in adult mice, including
408 classical V-shaped frequency response areas [22, 29-32]. A representative example from a
409 responsive neuron of an adolescent expert and one from an adult expert are shown in **Figure**
410 **7b,c**. To assess differences in pure-tone responses between auditory subregions, we quantified
411 the firing properties of AUDp versus AUDv (**Supplementary Table 8**; TEa and AUDd were
412 excluded since fewer than 20 units responded to the FRA protocol after task engagement, see
413 **Supplementary Table 7**). We found that the latency to peak was slower in adolescent AUDp.
414 The FWHM and minimal latency were slower in the adolescents in both AUDp and AUDv. In
415 the adolescent AUDv, additional parameters (like spontaneous- & evoked-FR, trial-

416 responsiveness, life time sparseness), were different than those in adults. There were no
417 significant differences between AUDp and AUDv within age-groups (**Supplementary Table**
418 **8**).

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

435

436 # Discussion

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

445

446 *Adolescence behavior – perceptual and cognitive differences*

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

459 Another manifestation of weaker cognitive control mechanisms is the noisy behavioral
460 performance of adolescents (**Fig. 2h-j**). While this noise may reflect higher plasticity and
461 potentially serve cognitive flexibility[6, 40], it poses a disadvantage for the precision and
462 consistency required in our task. This noise may also arise from changes in pubertal hormones,

463 which we did not manipulate[41]. Notably, few studies have explicitly characterized auditory
464 discrimination of adolescents in rodents; the majority of these have been carried out in gerbils
465 using an amplitude modulation (AM) detection task [4, 26, 42]. Despite the different task, many
466 of our age-related findings are aligned with the findings in gerbils, including lower performance
467 in adolescents, and greater variance in adolescent behavior[4]. However, some findings differ.
468 For example, the lick bias which was different in our data, was similar between young and adult
469 gerbils in an AM modulation detection task. This difference may arise from differences in task
470 design, as well as the severity of the punishment used in the task (here, we used a milder
471 punishment). Here, our data support a view that cognitive factors, and to a lesser extent
472 perception of pure tones *per se*, are the primary contributors to the poorer performance observed
473 in adolescent mice when performing a pure-tone sound discrimination task.
474
475 *Representation of stimulus and choice in ACx still undergoes maturation in adolescence*
476 Motivated by the abovementioned behavioral differences, we tested how single neurons in ACx
477 represented sounds (by comparing Hits vs FA) and choices (by comparing FA to CR). While
478 cortical representations of sounds in ACx could be the basis of perceptual constraints, distinct
479 processing of choices would reflect that more cognitive mechanisms are involved. We found
480 that cortical representations of sounds and choices in mice engaged in the task were still
481 immature as compared to adults, e.g., responses in adolescents were more delayed and less
482 informative (**Figs. 4-6**). These differences can be a result of age-related differences, and/or the
483 result of a combined 'age x learning' effect.
484 We found only modest age-related differences in spiking activity of layer 5/6 neurons
485 between adolescents and adults (**Fig. 7**). This limited difference is perhaps not surprising
486 because the CP for tonal organization in ACx is already complete by the age at which we tested
487 the animals (i.e., on P37-P42 [9, 13]). In contrast, more pronounced differences emerged from
488 the interaction of 'age x learning', specifically in mice that learned the task and were actively
489 engaged in the behavior (**Fig. 4f; Fig. 6d-g**). This interaction is evident in the lower single-
490 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

519 neurons. This is consistent with our recent work showing that indeed the same neurons in ACx
520 respond to the same stimulus differently, depending on if the mouse is engaged in
521 discriminating an easy sound pair or a hard sound pair[46]. The shallower slopes in adolescence
522 suggest a less efficient decoding scheme that may arise from different populations of neurons
523 encoding task-related information.

524 Our optogenetic silencing of ACx in expert adult mice resulted in a marked suppression
525 of licking behavior (**Fig. S3-2**), consistent with the idea that ACx contributes to task
526 performance. However, the strength of the observed effect raises the possibility that the
527 manipulation may have impacted broader aspects of behavior—such as arousal, motivation, or
528 motor readiness—rather than selectively disrupting auditory discrimination. Recent work [61]
529 suggests that ACx plays a critical role during learning by supporting higher-order computations
530 but may become less essential for task execution once mice reach expert performance. These
531 findings, together with previous reports [19, 20], indicate that the necessity of ACx for auditory-
532 guided behavior may depend on task structure, training stage, and the exact nature of the
533 manipulation. The robust decrease in lick rate and performance observed in our experiment
534 may reflect the intensity, specificity, and timing of the silencing. The use of GtACR2, a highly
535 potent silencer, may have further amplified the effect. We cannot rule out the possibility that
536 silencing the output from ACx also disrupted the flow of information to downstream areas,
537 thereby exerting a stronger influence on perception and behavior. Although the injection was
538 targeted to the ACx, off-target effects are also likely to have contributed. Future experiments
539 using graded or temporally specific perturbations will be useful to more define more precisely
540 the contribution of ACx across learning and development.

541

542 *Learning in the adolescent brain*

543 Cortical representations of pure tones have a critical period (CP) between P12-P15[13, 16, 17].
544 Indeed, we chose to study animals at P37-P42 on a pure tone discrimination task precisely to
545 avoid the hypersensitivity associated with this CP. But this may be unavoidable, as recent work
546 has shown that ACx also has a later CP for more complex sounds features[16, 17]. Those late

547 CP's overlap with our training schedule and the underlying biological manifestation of this CP
548 is expected to affect learning related plasticity in general. The underlying mechanism of the late
549 CP to complex sounds has been associated with dynamic changes of inhibitory circuits[62, 63].
550 Thus, immature inhibitory circuits remain sensitive in adolescents, and could affect how pure
551 tones are represented, despite the closure of the classic tonal CP. Notably, sound discrimination
552 learning changes tonal representations even in adult mice, well after the closure of CP's for
553 both simple and complex sounds [18, 24, 46]. We hypothesize, therefore, that the differences
554 in performance on more difficult discriminations between adolescents and adults are shaped
555 primarily by age-specific learning-related plasticity, rather than by differences in pure tone
556 representation per se.

557 Taken together, our findings show that cortical responses in adolescents differ from
558 those in adults, despite both groups learning the same pure tone discrimination task under
559 identical training protocols. Although the CPs for pure tones are considered closed by this age,
560 learning to discriminate between tones still varies across developmental stages. It remains
561 unclear how later CPs, such as those shaping responses to amplitude modulation or frequency-
562 modulated sweeps through inhibitory circuits, contribute to these differences. Likewise, the role
563 of ongoing maturation in cognitive control, including the development of feedback pathways,
564 warrants further investigation. Nevertheless, it is evident that developmental processes in the
565 adolescent brain continue to influence how even simple sounds are encoded through learning.
566 The adolescent-specific plasticity observed here directly reflects on how perceptual and
567 cognitive features are integrated within the auditory cortex.
568

Acknowledgments:

We thank the Mizrahi lab for comments on the manuscript and Linda Wilbrecht, 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.

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

Declaration of Interests:

The authors declare no competing interests.

590 **Supplemental information:**

591 S1-1: *Behavioral criteria of auditory learning,* related to Figure 1.

592 S3-1: *Auditory Cortex is necessary for task execution in adult mice,* related to Figure 3.

593 S3-2: *Verification of GtACR2 expression,* related to Figure 3.

594 S3-3: *Adolescent and adult mice performed similarly throughout recordings as well as*

595 *between the head-fixed configuration and the Educage,* related to Figure 3.

596 S4-1: *Probe reconstruction and activity profile across ACx regions,* related to Figure 4.

597 S4-2: *The neuronal discriminability of stimulus- and choice related activity is similar across*

598 *auditory sub-regions,* related to Figure 4.

599 S5-1: *Decoding accuracy of the first 200ms after the response window,* related to Figure 5.

600 S6-1: *Behavioral performance of novice mice,* related to Figure 6.

601 S6-2*: The neuronal discriminability of Easy and Hard Go and No-Go are distributed similar*

602 *across auditory sub-regions in adolescent and adult novice and expert mice,* related to Figure

603 6

604 S7-1: *Learning related changes in neuronal tuning properties in primary and secondary*

605 *auditory cortex*, related to Figure 7.

606 Supplementary Table 1: *Adolescent and adult mice exhibit different lick behavior in the task,*

607 related to Figure 3.

608 Supplementary Table 2: *Neuronal statistics in expert and novice recordings during task*

609 *engagement,* related to Figure 4 and Figure 6.

610 Supplementary Table 3: *Adolescent and adult expert mice have distinct firing properties in*

611 *different sub-regions,* related to Figure 4.

612 Supplementary Table 4: *Adolescent and adult firing properties are distinct between different*

613 *sub-regions,* related to Figure 4.

614 Supplementary Table 5: *Novice, adolescent and adult mice have distinct firing properties in*

615 *different sub-regions*, related to Figure 6.

616 Supplementary Table 6: *Adolescent and adult firing properties of novice mice are distinct*

617 *between different sub-regions,* related to Figure 6.

Supplementary Table 7: *Neuronal statistics during passive FRA protocol in expert and novice mice,* related to Figure 7.

Supplementary Table 8: *Adolescent and adult mice have distinct firing properties across different sub-regions of ACx—passive listening*, related to Figure 7.

Supplementary Table 9*: Overview of datasets analyzed, related to all figures.*

626 # Methods

627 ***Animals***

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

634

635 ***Surgical procedures***

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

638

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

644

645 *Head-bar Implantation & Craniotomy.* 48h before the start behavioral training on the head-
646 fixed recording setup, mice were anesthetized with 2% Isoflurane, the body temperature
647 maintained at 37 ° C and the eyes were covered with 5% Chloramphenicol ointment to prevent
648 from drying. Before removal of the scalp, we applied 4% Lidocaine above the skin and skull
649 area. Afterwards, we glued a custom-made titanium bar at the back of the mouse skull using
650 dental cement (Meta-bond). To enable acute recordings in ACx, we performed a craniotomy on
651 the left-hemisphere at 2.5 mm posterior and 4.0 lateral to the bregma. The craniotomy was
652 protected by a pool of dental cement and covered with a silicon elastomer (WPI; Kwik-Cast

653 catalog #KWIK CAST). After the craniotomy, mice were injected with 0.04mg/g of 10%
654 Meloxicam solution and 0.2 ml Saline. During the 48 hours of recovery, mice received two
655 further doses of 0.04mg/g of 10% Meloxicam solution. 24h before recordings, mice were again
656 anesthetized to remove regrowth of the dura-mater during training. The area around the
657 craniotomy was cleaned with Hydrogen-peroxide (3%), the silicon elastomer was replaced, and
658 another dose of 10% Meloxicam was injected.
659
660 *Virus Injection & Cannula Implantation.* The surgical procedure was similar to the *head-bar*
661 *implantation & craniotomy*. The differences in the procedure are outlined below. Bilateral
662 craniotomies were performed at 2.5 mm posterior and 4.3 mm lateral to the bregma. We injected
663 200nl AAV5-CAMKII-GtACR2-FRED_kv_2.1, or 200nl AAV9-CAMKII-dTomato. Both
664 viruses were produced at the ELSC virus-core facility. We injected the virus at a depth of 1.1
665 mm at 0°. Afterwards, we inserted a 0.2 mm diameter optical fiber with an attached cannula
666 (CFML1202; Thorlabs) at a depth of 0.9 mm at 0°, bilaterally. We chronically fixed the fiber
667 position on the skull using dental cement (Meta bond).
668
669 ***Auditory Stimuli***
670 Sound stimuli were delivered from a calibrated free-field speaker (ES1 SN:4568) using a
671 multifunction processor (ED1; Tucker-Davis Technologies). The speakers were calibrated with
672 a free-field microphone (Type 4939, Bruel and Kjaer). The stimuli were comprised of pure
673 tones, sampled at 500 kHz, between 7.07 to 14 kHz for behavioral training and 4-40 kHz for
674 passive listening. All stimuli were played at sound pressure levels of 72 dB SPL for behavioral
675 training, and 72-, 62-, 52-, 42-, and 32-dB SPLs for passive listening. All tones were played in
676 randomized order.
677
678 ***Behavior***
679 *Automated home-cage training (the 'Educage').* For a detailed description of the EduCage
680 behavioral platform, see [18]. After RFID- implantation, groups of 5 mice were placed in their

681 home-cage that was connected to a behavioral chamber called the Educage (Fig. 1a). Mice were
682 provided with food *ad libitum*. Water access was restricted to the Educage chamber. Mice could
683 access the Educage freely to retrieve water mixed with 5% sucrose. The behavioral training was
684 controlled by a custom LabView 2019 (National Instruments) program running on BNC-2110
685 (National Instruments) and FPGA (MyRIO – National Instruments) to register licks, RFIDs
686 numbers, and deliver stimuli and reinforcements. Every time a mouse entered the water port its
687 RFID was automatically recognized, a trial initiated, and an auditory stimulus played from a
688 speaker vertically positioned above the behavioral chamber. We trained mice gradually for
689 three weeks on a Go/No-Go paradigm of pure tones, spaced around a category boundary of 10
690 kHz. This paradigm enabled us to efficiently compare between adolescent and adult behavior,
691 since learning was restricted from post-weaning (P20) to adolescence (P37) (correspondingly
692 adults were trained from P60-P77). Mice were trained in four separate training phases (Fig.
693 1B). For the first 24h (P20-P21 in the adolescent group of mice and P60-P61 in adult mice)
694 mice were habituated to the behavioral chamber and could approach the port freely to retrieve
695 water after licking the waterspout (0.0015 ml per trial). During the following 48h (P21-P23 in
696 the adolescent group of mice and P61-P63 in adult mice) mice underwent a tone-association
697 phase. All tones played were 300ms long (sampled at 500 kHz). Tone association was restricted
698 to the lowest Go tone (7.07 kHz at 100% probability). Mice received a water-reward in response
699 to licking during the first 2 seconds after tone-offset. This was followed by the first tone-
700 discrimination stage, which contained a No-Go tone (14.14 kHz at 45% probability) at 1 octave
701 away from the Go tone. We analyzed four response outcomes—Hits, Misses, Correct
702 Rejections, False Alarms— in the 2 seconds response window after the tone-offset. 'Hits' were
703 counted after 5 successful licks to the Go-tone and followed by a water-reward. 'Misses' were
704 counted when licks to the Go-tone trials did not pass the lick threshold. Correspondingly,
705 'Correct Reject' (CR) trials, were all No-Go trials below the lick threshold. CR trials were not
706 reinforced. 'False Alarms' (FA) were counted when 5 or more licks were registered after the
707 No-Go. FA trials were followed by a white-noise punishment (5-20 kHz at 72 dB SPLs, 2
708 seconds) and an inter trial delay of 5 seconds. In addition, mice were exposed to 5 probe trials

709 (tones whose trial outcome that were neither rewarded nor punished at 8.49 kHz, 9.567 kHz,
710 10 kHz, 10.44 kHz, and 11.89 kHz; played at 2% per tone). Mice trained on the 1 octave tone
711 discrimination for 1 week (P23-P30 in adolescent mice and P63-P70 in adult mice). Next, we
712 added a second pair of tone - a Go/No-Go tone pair spaced 0.25 octaves apart (9.17 kHz and
713 10.95 kHz) for another week of training. All tones were played in randomized order and trial
714 difficulties of the easy and the hard task were intermingled. The experiment was terminated at
715 age P37 for adolescent and P77 for adult mice.
716
717 *Head-fixed training*. We used a head-fixed paradigm that enabled us to efficiently train
718 adolescent and adult mice on the same task and record from engaged mice after learning. We
719 trained mice on the same Go/No-Go stimulus setup of the Educage described above with few
720 differences which are outlined below.
721 The head-fixed behavioral training was controlled by a custom MATLAB 2023B
722 (Mathworks) program running on BNC-2110 (National Instruments) to register licks and
723 deliver stimuli and reinforcements. Mice were water-restricted for 24h prior to training, and
724 their weight monitored daily. While training on the setup, mice received water supplemented
725 with 5% sucrose, depending on their weight (0.125 ml per g of body weight). Additional water,
726 in the form of Hydrogel (ClearH2O) was given after daily training if mice did not consume the
727 required amount during training, or their body weight dropped below 85% of the pre-restriction
728 weight. Mice were gradually habituated to the setup by head-fixation and received free water
729 after licking, during the first training day (P23 in adolescent mice and P63in adult mice).
730 Licking was measured with an optical lick-meter (Sanworks). During the second training day
731 mice were associated with a tone. All tones of the task were 100ms long (sampled at 500 kHz).
732 Mice received a water reward in response to licking in the first 2 seconds after tone offset.
733 Reward-reinforcement was delayed to 0.5 seconds after the tone offset. To break the regularity
734 of the trial sequences, we applied a dynamic inter-trial interval of 6-8 seconds. New trials were
735 initiated if mice did not lick the spout at least 2 seconds prior to the upcoming trial. Mice were
736 familiarized with the tone for at least two training days (P24-26 in adolescent mice and P64-66

737 in adult mice). Then, and when the lick rate was above 80%, mice proceeded to the first tone
738 discrimination stage. We gradually increased the probability of No-Go trials, depending on the
739 lick rate of the mouse. We added the 0.25 octave Go/No-Go tone pair, after a maximal of seven
740 training days on the 1 octave tone pair (P26-33 in adolescent mice and P66-73 in adult mice),
741 or if the behavioral performance exceeded a threshold of d' = 1, as calculated using the signal
742 detection metric. Similar to the Educage, all tones were played in randomized order and trial
743 difficulties of the easy and the hard task were intermingled. Mice were trained until P37 in the
744 adolescent group and P77 in the adult group.
745
746 ***Optogenetic Manipulation***
747 Adult mice (P60) were injected with AAV5-CAMKII-GtACR2-FRED_kv_2.1 (n = 3) and
748 implanted with cannulas prior to training as described above. Mice underwent optogenetic
749 stimulation after training on the head-fixed task by attaching a ferrule patch cable (M79L01;
750 Thorlabs) to the implanted cannula using a mating sleeve (ADAF1; Thorlabs). The ferrule cable
751 end (SMA connector) was connected to an LED driver setup (LEDD1B; Thorlabs) that enabled
752 us to send precise outputs at 476 nm (blue light; 5mW) bilaterally. To test the effect of ACx
753 inhibition on task performance, we timed the LED output to 50ms before tone onset until 50ms
754 after tone-offset. The LED consisted of pulses given at 10 Hz for 200ms
755 Behavioral sessions of optogenetic manipulation were identical to the Go/No-Go
756 paradigm of previous training sessions. We inhibited ACx in 50% of the trials in randomized
757 order. Optogenetic manipulations were repeated throughout multiple sessions of expert
758 performance. As a control, we repeated the same experiment in mice injected with AAV9-
759 CAMKII-dTomato (n = 3). After completion of the experiment, animals were deeply
760 anesthetized with Ketamine-and medetomidine (0.008 g/kg, and 0.65 mg/kg, respectively) and
761 then perfused with 4% Paraformaldehyde. After the perfusion, the brain was extracted and
762 preserved. We then sectioned the ACx using a cryostat (Leica) into 0.05 mm thick coronal
763 slices. Afterwards brain slices were washed one time under 1% Phosphate-buffered saline
764 (PBS), and a second time with 1% PBS plus 0.4% Triton. Finally, slides were mounted and

765 stained with DAPI (4%) and imaged using a wide-field fluorescent microscope (Olympus Life
766 Science, Olympus IX83), to reconstruct the probe position offline.
767
768 ***Extracellular Recordings***
769 *Recording setup.* Before the onset of the recording session, we removed the silicon elastomer
770 and placed an external reference electrode (Ag/AgCl wire) on the skull of the right hemisphere.
771 All recordings were performed using Neuropixels 1.0 (Npx; IMEC, phase 3A), together with a
772 base-station connected to a chassis (IMEC; NI PXIe -1071, National Instruments). Probes were
773 mounted to a custom-made steel rod and connected to the ground. Before penetration, probes
774 were covered with a fluorescent dye (Dil (Invitrogen catalog #V22885- red, or Dio (Invitrogen
775 catalog #V22886 – yellow), to reconstruct penetration sites after the recording.
776 Then, Npx probes were inserted into the left ACx at 30° to a depth of 3.85 mm (this approach
777 limited our recording to deep layers of the ACx). The skull surface was submerged in saline
778 and the probe allowed to settle for 10 min. We sampled recordings at 20 kHz, with action
779 potential band filtered to contain 0.3- to 10kHz frequencies. Action potential band gain was set
780 to 500. Out of the 960 available sites on the 1 cm shank of the Npx probe [68], we acquired the
781 384 lowest recording shanks in a staggered configuration. We used common-average
782 referencing to process the voltage-traces.
783
784 *Recording Schedule.* We performed recordings of both novice and expert adolescent and adult
785 mice. Novice recordings were performed after tone-association (see *head-fixed training;*
786 adolescent n=3; adult n = 3). Expert recordings were performed after learning (adolescent n=
787 5; adult n = 6). Both novice and expert recording include the easy (1 octave) and hard (0.25
788 octave) tone pair. Both recordings were performed after P37 for adolescent mice and after P77
789 for adult mice. Tone probabilities and protocol parameters were identical to the *head-fixed*
790 *training*. Each mouse was recorded multiple times (novice recordings adolescents n = 6; novice
791 recording adults n =6; expert recordings in adolescents n = 13; expert recordings in adults n =
792 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 μV, SNR > 3), physiologically plausible refractory period (ISI > 4ms), 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' = the inverse normal distribution of the z-scored hit rate) – 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:

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

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-scored (hit rate) + z-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:

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

873 r_i = corresponds to the FR of each learned frequency and n equals the number of learned tones.
874 Values of S near 0 correspond to a dense FR and values near 1 to a sparse code [71].
875
876 *Single neuron analysis during task performance.* We investigated the FR of excitatory auditory
877 responses per trial from 200ms preceding the stimulus up to 600ms after the tone onset. To test
878 how well single neurons discriminate between trial outcomes, as well as task difficulty, we
879 calculated task-related activity per trial using receiver operating characteristics (ROC) and
880 calculated the area under the curve (AUC) across a running window of 25ms in bins of 50ms
881 (Fig. 4, Fig. 6). Stimulus-related activity was defined as the hit compared to the false alarm
882 trials, separately for the 1 octave and 0.25 octave tone pair. Stimulus-related activity of miss
883 compared to correct trials were not analyzed due to the lack of miss trials in expert performance.
884 In addition, we also tested the average (mean of 1 octave and 0.25 octave tone pair) choice
885 discrimination, defined as the difference between false alarm and correct reject trials (Fig. 4).
886 We sampled 20 trials per trial outcome and repeated the ROC encoding 100 times. Afterwards,
887 we adjusted the average AUC values of all iterations of each neuron to its baseline AUC before
888 the tone onset ((AUC – baseline AUC) + 0.5). AUC values close to 0.5 indicate a low neuronal
889 discriminability, and values up to 1 indicate high neuronal discriminability [72]. To test
890 significant discrimination, we repeated the ROC encoding with shuffled trial identities (20
891 randomly shuffled trials per trial outcomes across 100 iterations). Significant discriminability
892 of a neuron was defined as the AUC time bins that exceed the mean ±3 std of the shuffled
893 distribution [64]. Adolescent and adult neuronal discriminability were compared using the
894 onset-latency of discrimination of a neuron (first timepoint > mean ±3 std of the shuffled
895 distribution), the maximal AUC of the running window and the duration of discriminability
896 (time duration > mean ±3 std of the shuffled distribution;). To compare novice and expert
897 performance we repeated the AUC encoding and compared the general discrimination of 1
898 octave and 0.25 octave Go and No-Go discrimination of all trial outcomes (Fig. 6). To evaluate
899 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:

$$S = \frac{\sigma^2_{between}}{\sigma^2_{within}} = \max_{w} \frac{\left(w(\mu_1 - \mu_2)\right)^2}{w^T(\Sigma_1 + \Sigma_2)w}. \qquad \text{(Equation 3)}$$

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

927 and tasks. Data were standardized per neuron before computing variance per stimulus (Fig. 5e).
928 To visualize the results, we projected the population representations to the space spanned by
929 the two LDA decoding dimensions for the easy and hard tasks. We averaged the mean and
930 covariance matrix across sessions for each stimulus and age group, representing the
931 distributions with corresponding ellipses (Fig. 5f).
932
933 *Single neuron analysis during passive listening.* We extracted the frequency-response areas
934 (FRAs) based on the pure-tone responses (from tone onset up to 50ms after tone-offset) of
935 auditory responsive neurons. The bandwidth of excited units was computed for significant
936 excitatory responses of adjacent frequencies at 62 dB SPL and subtracted by the number of
937 expected false-positive responses. Afterwards the sum of significant frequency responses per
938 neuron was multiplied by the octave distance between every frequency to receive an octave-
939 based bandwidth measure. We then calculated the population sparseness as the fraction of
940 significant excited responses in adolescent and adult novice and expert mice. To test the
941 neuronal discriminability to pure tone responses, we calculated the pairwise d' based on the FR
942 per trial in all frequencies at 62 dB SPL [22, 66]. For two given frequencies p and q, d' values
943 were calculated in the following way:

944 $$\frac{d(\mu_{p,}\mu_q)}{\mu(\sqrt{\sum_n^{i=1}(\widehat{p_i}-\widehat{q_i})^2})} \qquad \text{(Equation 4)}$$

945
946 Where d represents the distance between the mean FR (μ) of frequency p and q, as vectors with
947 n entries representing the mean signal (averaged over individual trials) in n-dimensional space
948 for frequency p/q. The average distance (d) is then divided by the mean of the inner Euclidian
949 distance ($\sqrt{\sum_n^{i=1}(\widehat{p_i}-\widehat{q_i})}$) between each single trial (Fig. 7h).
950
951 *Statistical Analysis.* Statistical comparisons were performed in custom-written codes in
952 MATLAB 2023b (MathWorks) and Python 3.12. We assessed if the data was normally

953 distributed using a Kolmogorov-Smirnov test. Normally distributed data was tested using a
954 paired (within group comparison) or independent (between group comparison) t-test and
955 presented with mean ± STE. Non-normally distributed data was tested using a Wilcoxon sign
956 rank (within group comparison), or rank-sum (between group comparison) test. Multiple-
957 samples were tested with ANOVA (parametric data) and Kruskal-Wallis tests (non-parametric
958 data). All tests corrected after multiple comparisons (Bonferroni-correction for two-samples
959 and Tukey-Kramer for multiple samples). In addition, we applied linear-mixed-effect models
960 (LME) to account for hierarchical data structures and variability within groups of co-housed
961 mice, as well as repeated measurements of mice and recordings, as random effects. Fixed
962 effects, and pairwise interaction effects were adjusted with post-hoc tests and multiple
963 comparisons. For all statistical tests significant differences were defined as p-values below 0.05.
964

Figure Legends

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). **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'>=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 ± 0.1791; # trials = 8317 ± 712) and adult mice at P70 (black; n=15, d'=2.1986 ± 0.2441; # trials = 13229 ± 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 (Δd') 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; Δd' 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; Δd' 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 ± 0.1783; # trials=4163 ± 297; adult-black; n=15, d'=1.8342 ± 0.1743; # trials=4102±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; Δd' between adult and adolescent mice: z = -1.9495; p = 0.0512, two-sample Wilcoxon rank sum test).

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.

Figure 3: *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.

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). **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 -200ms to +600ms 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 -200ms to 600ms 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 (+- 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).

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, pared t-test; hard task: p = 0.0100). **C.** Decoding accuracy over a time window from -0.5s to 10s (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. **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 +-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 (+-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.

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. **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).

Main Tables

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') ~ age + sex + (1|cage ID).

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).

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 10). 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).

Supplementary Figure Legends

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 *i*nter 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 = 6.8241e-05, 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). **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 = 2.4414e-04; adult mice: signed rank: 105, p = 1.2207e-04, 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 μm bins) across AUDd, AUDp, AUDv and TEa (3850 maximal depth μm). **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, AUDv, 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.

Supplementary Tables

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 * dprime + (1|Mouse ID) + (1|Recording ID).

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).

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).

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).

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).

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).

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.

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).

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.

Code availability

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

References

1. Hoskins D. Consequences of Parenting on Adolescent Outcomes. Societies. 2014;4(3):506-31. doi: 10.3390/soc4030506.
2. Hawkins JD, Catalano RF, Miller JY. Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention. Psychol Bull. 1992;112(1):64-105. doi: 10.1037/0033-2909.112.1.64. PubMed PMID: 1529040.
3. Fuhrmann D, Knoll LJ, Blakemore SJ. Adolescence as a Sensitive Period of Brain Development. Trends Cogn Sci. 2015;19(10):558-66. doi: 10.1016/j.tics.2015.07.008. PubMed PMID: 26419496.
4. Caras ML, Sanes DH. Neural Variability Limits Adolescent Skill Learning. J Neurosci. 2019;39(15):2889-902. Epub 20190212. doi: 10.1523/JNEUROSCI.2878-18.2019. PubMed PMID: 30755494; PubMed Central PMCID: PMCPMC6462443.
5. Huyck JJ, Wright BA. Late maturation of auditory perceptual learning. Dev Sci. 2011;14(3):614-21. Epub 20101111. doi: 10.1111/j.1467-7687.2010.01009.x. PubMed PMID: 21477199; PubMed Central PMCID: PMCPMC3107547.
6. Johnson C, Wilbrecht L. Juvenile mice show greater flexibility in multiple choice reversal learning than adults. Dev Cogn Neurosci. 2011;1(4):540-51. doi: 10.1016/j.dcn.2011.05.008. PubMed PMID: 21949556; PubMed Central PMCID: PMCPMC3177141.
7. Romer D. Adolescent risk taking, impulsivity, and brain development: implications for prevention. Dev Psychobiol. 2010;52(3):263-76. doi: 10.1002/dev.20442. PubMed PMID: 20175097; PubMed Central PMCID: PMCPMC3445337.
8. Wilbrecht L, Davidow JY. Goal-directed learning in adolescence: neurocognitive development and contextual influences. Nat Rev Neurosci. 2024;25(3):176-94. Epub 20240123. doi: 10.1038/s41583-023-00783-w. PubMed PMID: 38263216.
9. Hensch TK. Critical period plasticity in local cortical circuits. Nat Rev Neurosci. 2005;6(11):877-88. doi: 10.1038/nrn1787. PubMed PMID: 16261181.
10. Hensch TK. Critical period regulation. Annu Rev Neurosci. 2004;27:549-79. doi: 10.1146/annurev.neuro.27.070203.144327. PubMed PMID: 15217343.
11. Reh RK, Dias BG, Nelson CA, 3rd, Kaufer D, Werker JF, Kolb B, et al. Critical period regulation across multiple timescales. Proc Natl Acad Sci U S A. 2020;117(38):23242-51. Epub 20200605. doi: 10.1073/pnas.1820836117. PubMed PMID: 32503914; PubMed Central PMCID: PMCPMC7519216.
12. Kral A. Auditory critical periods: a review from system's perspective. Neuroscience. 2013;247:117-33. Epub 20130521. doi: 10.1016/j.neuroscience.2013.05.021. PubMed PMID: 23707979.
13. Barkat TR, Polley DB, Hensch TK. A critical period for auditory thalamocortical connectivity. Nat Neurosci. 2011;14(9):1189-94. Epub 20110731. doi: 10.1038/nn.2882. PubMed PMID: 21804538; PubMed Central PMCID: PMCPMC3419581.
14. Sun YJ, Wu GK, Liu BH, Li P, Zhou M, Xiao Z, et al. Fine-tuning of pre-balanced excitation and inhibition during auditory cortical development. Nature. 2010;465(7300):927-31. doi: 10.1038/nature09079. PubMed PMID: 20559386; PubMed Central PMCID: PMCPMC2909826.
15. Zhang LI, Bao S, Merzenich MM. Persistent and specific influences of early acoustic environments on primary auditory cortex. Nat Neurosci. 2001;4(11):1123-30. doi: 10.1038/nn745. PubMed PMID: 11687817.
16. Bhumika S, Nakamura M, Valerio P, Solyga M, Linden H, Barkat TR. A Late Critical Period for Frequency Modulated Sweeps in the Mouse Auditory System. Cereb Cortex. 2020;30(4):2586-99. doi: 10.1093/cercor/bhz262. PubMed PMID: 31800018; PubMed Central PMCID: PMCPMC7174992.

1461 17. Nakamura M, Valerio P, Bhumika S, Barkat TR. Sequential Organization of Critical
1462 Periods in the Mouse Auditory System. Cell Rep. 2020;32(8):108070. doi:
1463 10.1016/j.celrep.2020.108070. PubMed PMID: 32846128.
1464 18. Maor I, Shwartz-Ziv R, Feigin L, Elyada Y, Sompolinsky H, Mizrahi A. Neural Correlates
1465 of Learning Pure Tones or Natural Sounds in the Auditory Cortex. Front Neural Circuits.
1466 2019;13:82. Epub 20200128. doi: 10.3389/fncir.2019.00082. PubMed PMID: 32047424;
1467 PubMed Central PMCID: PMCPMC6997498.
1468 19. O'Sullivan C, Weible AP, Wehr M. Auditory Cortex Contributes to Discrimination of
1469 Pure Tones. eNeuro. 2019;6(5). Epub 20191016. doi: 10.1523/ENEURO.0340-19.2019.
1470 PubMed PMID: 31591138; PubMed Central PMCID: PMCPMC6795560.
1471 20. Ceballo S, Piwkowska Z, Bourg J, Daret A, Bathellier B. Targeted cortical manipulation
1472 of auditory perception. Neuron. 2019;104(6):1168-79. e5.
1473 21. Kanold PO, Nelken I, Polley DB. Local versus global scales of organization in auditory
1474 cortex. Trends Neurosci. 2014;37(9):502-10. Epub 20140704. doi:
1475 10.1016/j.tins.2014.06.003. PubMed PMID: 25002236; PubMed Central PMCID:
1476 PMCPMC4152386.
1477 22. Feigin L, Tasaka G, Maor I, Mizrahi A. Sparse Coding in Temporal Association Cortex
1478 Improves Complex Sound Discriminability. J Neurosci. 2021;41(33):7048-64. Epub 20210708.
1479 doi: 10.1523/JNEUROSCI.3167-20.2021. PubMed PMID: 34244361; PubMed Central PMCID:
1480 PMCPMC8372025.
1481 23. Bao S, Chang EF, Teng CL, Heiser MA, Merzenich MM. Emergent categorical
1482 representation of natural, complex sounds resulting from the early post-natal sound
1483 environment. Neuroscience. 2013;248:30-42. Epub 20130606. doi:
1484 10.1016/j.neuroscience.2013.05.056. PubMed PMID: 23747304; PubMed Central PMCID:
1485 PMCPMC3838508.
1486 24. Bao S, Chang EF, Woods J, Merzenich MM. Temporal plasticity in the primary
1487 auditory cortex induced by operant perceptual learning. Nat Neurosci. 2004;7(9):974-81.
1488 Epub 20040801. doi: 10.1038/nn1293. PubMed PMID: 15286790.
1489 25. Han YK, Kover H, Insanally MN, Semerdjian JH, Bao S. Early experience impairs
1490 perceptual discrimination. Nat Neurosci. 2007;10(9):1191-7. Epub 20070729. doi:
1491 10.1038/nn1941. PubMed PMID: 17660815.
1492 26. Anbuhl KL, Yao JD, Hotz RA, Mowery TM, Sanes DH. Auditory processing remains
1493 sensitive to environmental experience during adolescence in a rodent model. Nat Commun.
1494 2022;13(1):2872. Epub 20220524. doi: 10.1038/s41467-022-30455-9. PubMed PMID:
1495 35610222; PubMed Central PMCID: PMCPMC9130260.
1496 27. Jaramillo S, Zador AM. The auditory cortex mediates the perceptual effects of
1497 acoustic temporal expectation. Nat Neurosci. 2011;14(2):246-51. Epub 20101219. doi:
1498 10.1038/nn.2688. PubMed PMID: 21170056; PubMed Central PMCID: PMCPMC3152437.
1499 28. Kato HK, Gillet SN, Isaacson JS. Flexible Sensory Representations in Auditory Cortex
1500 Driven by Behavioral Relevance. Neuron. 2015;88(5):1027-39. Epub 20151112. doi:
1501 10.1016/j.neuron.2015.10.024. PubMed PMID: 26586181; PubMed Central PMCID:
1502 PMCPMC4670799.
1503 29. Rothschild G, Nelken I, Mizrahi A. Functional organization and population dynamics
1504 in the mouse primary auditory cortex. Nat Neurosci. 2010;13(3):353-60. Epub 20100131. doi:
1505 10.1038/nn.2484. PubMed PMID: 20118927.
1506 30. Cohen L, Mizrahi A. Plasticity during motherhood: changes in excitatory and
1507 inhibitory layer 2/3 neurons in auditory cortex. J Neurosci. 2015;35(4):1806-15. doi:
1508 10.1523/JNEUROSCI.1786-14.2015. PubMed PMID: 25632153; PubMed Central PMCID:
1509 PMCPMC6795264.

31. Cohen L, Rothschild G, Mizrahi A. Multisensory integration of natural odors and sounds in the auditory cortex. Neuron. 2011;72(2):357-69. doi: 10.1016/j.neuron.2011.08.019. PubMed PMID: 22017993.
32. Maor I, Shalev A, Mizrahi A. Distinct Spatiotemporal Response Properties of Excitatory Versus Inhibitory Neurons in the Mouse Auditory Cortex. Cereb Cortex. 2016;26(11):4242-52. doi: 10.1093/cercor/bhw266. PubMed PMID: 27600839; PubMed Central PMCID: PMCPMC5066836.
33. Sawyer SM, Afifi RA, Bearinger LH, Blakemore SJ, Dick B, Ezeh AC, et al. Adolescent Health 1 Adolescence: a foundation for future health. Lancet. 2012;379(9826):1630-40. doi: 10.1016/S0140-6736(12)60072-5. PubMed PMID: WOS:000303452600037.
34. Spear LP. Neurobehavioral changes in adolescence. Current Directions in Psychological Science. 2000;9(4):111-4. doi: Doi 10.1111/1467-8721.00072. PubMed PMID: WOS:000088613200001.
35. Dahl RE, Allen NB, Wilbrecht L, Suleiman AB. Importance of investing in adolescence from a developmental science perspective. Nature. 2018;554(7693):441-50. doi: 10.1038/nature25770. PubMed PMID: 29469094.
36. Reynolds LM, Yetnikoff L, Pokinko M, Wodzinski M, Epelbaum JG, Lambert LC, et al. Early Adolescence is a Critical Period for the Maturation of Inhibitory Behavior. Cereb Cortex. 2019;29(9):3676-86. doi: 10.1093/cercor/bhy247. PubMed PMID: 30295713; PubMed Central PMCID: PMCPMC6686753.
37. Meyer HC, Bucci DJ. Age differences in appetitive Pavlovian conditioning and extinction in rats. Physiol Behav. 2016;167:354-62. Epub 20161011. doi: 10.1016/j.physbeh.2016.10.004. PubMed PMID: 27737779; PubMed Central PMCID: PMCPMC5159263.
38. Magis-Weinberg L, Custers R, Dumontheil I. Rewards Enhance Proactive and Reactive Control in Adolescence and Adulthood. Soc Cogn Affect Neurosci. 2019;14(11):1219-32. doi: 10.1093/scan/nsz093. PubMed PMID: 31820793; PubMed Central PMCID: PMCPMC7057287.
39. Cohen JR, Asarnow RF, Sabb FW, Bilder RM, Bookheimer SY, Knowlton BJ, et al. A unique adolescent response to reward prediction errors. Nat Neurosci. 2010;13(6):669-71. Epub 20100516. doi: 10.1038/nn.2558. PubMed PMID: 20473290; PubMed Central PMCID: PMCPMC2876211.
40. Crone EA, Dahl RE. Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat Rev Neurosci. 2012;13(9):636-50. doi: 10.1038/nrn3313. PubMed PMID: 22903221.
41. Kunkhyen T, Perez E, Bass M, Coyne A, Baum MJ, Cherry JA. Gonadal hormones, but not sex, affect the acquisition and maintenance of a Go/No-Go odor discrimination task in mice. Horm Behav. 2018;100:12-9. Epub 20180307. doi: 10.1016/j.yhbeh.2018.02.009. PubMed PMID: 29481807; PubMed Central PMCID: PMCPMC5962265.
42. Buran BN, Sarro EC, Manno FA, Kang R, Caras ML, Sanes DH. A sensitive period for the impact of hearing loss on auditory perception. J Neurosci. 2014;34(6):2276-84. doi: 10.1523/JNEUROSCI.0647-13.2014. PubMed PMID: 24501366; PubMed Central PMCID: PMCPMC3913871.
43. Lee T-Y, Weissenberger Y, King AJ, Dahmen JC. Midbrain encodes sound detection behavior without auditory cortex. Elife. 2024;12:RP89950.
44. Kato HK, Asinof SK, Isaacson JS. Network-level control of frequency tuning in auditory cortex. Neuron. 2017;95(2):412-23. e4.
45. Talwar SK, Musial PG, Gerstein GL. Role of mammalian auditory cortex in the perception of elementary sound properties. Journal of Neurophysiology. 2001;85(6):2350-8. doi: DOI 10.1152/jn.2001.85.6.2350. PubMed PMID: WOS:000169280200004.

46. Haimson B, Gilday OD, Lavi-Rudel A, Sagi H, Lottem E, Mizrahi A. Single neuron responses to perceptual difficulty in the mouse auditory cortex. Sci Adv. 2024;10(33):eadp9816. Epub 20240814. doi: 10.1126/sciadv.adp9816. PubMed PMID: 39141740; PubMed Central PMCID: PMCPMC11323952.
47. Ahissar M, Hochstein S. The reverse hierarchy theory of visual perceptual learning. Trends Cogn Sci. 2004;8(10):457-64. doi: 10.1016/j.tics.2004.08.011. PubMed PMID: 15450510.
48. Ahissar M, Hochstein S. Task difficulty and the specificity of perceptual learning. Nature. 1997;387(6631):401-6. doi: 10.1038/387401a0. PubMed PMID: 9163425.
49. Zagha E. Shaping the cortical landscape: functions and mechanisms of top-down cortical feedback pathways. Frontiers in Systems Neuroscience. 2020;14:33.
50. Tasaka GI, Feigin L, Maor I, Groysman M, DeNardo LA, Schiavo JK, et al. The Temporal Association Cortex Plays a Key Role in Auditory-Driven Maternal Plasticity. Neuron. 2020;107(3):566-79 e7. Epub 20200529. doi: 10.1016/j.neuron.2020.05.004. PubMed PMID: 32473095; PubMed Central PMCID: PMCPMC8725613.
51. Tasaka GI, Maggi C, Taha E, Mizrahi A. The local and long-range input landscape of inhibitory neurons in mouse auditory cortex. J Comp Neurol. 2023;531(4):502-14. Epub 20221201. doi: 10.1002/cne.25437. PubMed PMID: 36453284; PubMed Central PMCID: PMCPMC10107844.
52. Delevich K, Klinger M, Okada NJ, Wilbrecht L. Coming of age in the frontal cortex: The role of puberty in cortical maturation. Semin Cell Dev Biol. 2021;118:64-72. Epub 20210510. doi: 10.1016/j.semcdb.2021.04.021. PubMed PMID: 33985902.
53. Konstantoudaki X, Chalkiadaki K, Vasileiou E, Kalemaki K, Karagogeos D, Sidiropoulou K. Prefrontal cortical-specific differences in behavior and synaptic plasticity between adolescent and adult mice. J Neurophysiol. 2018;119(3):822-33. Epub 20171122. doi: 10.1152/jn.00189.2017. PubMed PMID: 29167323.
54. Johnson CM, Loucks FA, Peckler H, Thomas AW, Janak PH, Wilbrecht L. Long-range orbitofrontal and amygdala axons show divergent patterns of maturation in the frontal cortex across adolescence. Dev Cogn Neurosci. 2016;18:113-20. Epub 20160204. doi: 10.1016/j.dcn.2016.01.005. PubMed PMID: 26896859; PubMed Central PMCID: PMCPMC5283395.
55. Nabel EM, Garkun Y, Koike H, Sadahiro M, Liang A, Norman KJ, et al. Adolescent frontal top-down neurons receive heightened local drive to establish adult attentional behavior in mice. Nat Commun. 2020;11(1):3983. Epub 20200807. doi: 10.1038/s41467-020-17787-0. PubMed PMID: 32770078; PubMed Central PMCID: PMCPMC7414856.
56. Druckmann S, Chklovskii DB. Neuronal circuits underlying persistent representations despite time varying activity. Curr Biol. 2012;22(22):2095-103. Epub 20121018. doi: 10.1016/j.cub.2012.08.058. PubMed PMID: 23084992; PubMed Central PMCID: PMCPMC3543774.
57. Kang B, Druckmann S. Approaches to inferring multi-regional interactions from simultaneous population recordings: Inferring multi-regional interactions from simultaneous population recordings. Curr Opin Neurobiol. 2020;65:108-19. Epub 20201120. doi: 10.1016/j.conb.2020.10.004. PubMed PMID: 33227602; PubMed Central PMCID: PMCPMC7853322.
58. Li N, Daie K, Svoboda K, Druckmann S. Robust neuronal dynamics in premotor cortex during motor planning. Nature. 2016;532(7600):459-64.
59. Wei Z, Inagaki H, Li N, Svoboda K, Druckmann S. An orderly single-trial organization of population dynamics in premotor cortex predicts behavioral variability. Nat Commun. 2019;10(1):216. Epub 20190115. doi: 10.1038/s41467-018-08141-6. PubMed PMID: 30644387; PubMed Central PMCID: PMCPMC6333792.

60. Wei Z, Lin BJ, Chen TW, Daie K, Svoboda K, Druckmann S. A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology. PLoS Comput Biol. 2020;16(9):e1008198. Epub 20200915. doi: 10.1371/journal.pcbi.1008198. PubMed PMID: 32931495; PubMed Central PMCID: PMCPMC7518847.
61. Drieu C, Zhu Z, Wang Z, Fuller K, Wang A, Elnozahy S, et al. Rapid emergence of latent knowledge in the sensory cortex drives learning. Nature. 2025:1-11.
62. Takesian AE, Kotak VC, Sanes DH. Age-dependent effect of hearing loss on cortical inhibitory synapse function. J Neurophysiol. 2012;107(3):937-47. Epub 20111116. doi: 10.1152/jn.00515.2011. PubMed PMID: 22090457; PubMed Central PMCID: PMCPMC3289466.
63. Takesian AE, Kotak VC, Sanes DH. Presynaptic GABA(B) receptors regulate experience-dependent development of inhibitory short-term plasticity. J Neurosci. 2010;30(7):2716-27. doi: 10.1523/JNEUROSCI.3903-09.2010. PubMed PMID: 20164356; PubMed Central PMCID: PMCPMC3842473.
64. Gilad A, Maor I, Mizrahi A. Learning-related population dynamics in the auditory thalamus. Elife. 2020;9:e56307. Epub 20200708. doi: 10.7554/eLife.56307. PubMed PMID: 32639231; PubMed Central PMCID: PMCPMC7371423.
65. Gilday OD, Praegel B, Maor I, Cohen T, Nelken I, Mizrahi A. Surround suppression in mouse auditory cortex underlies auditory edge detection. PLoS Comput Biol. 2023;19(1):e1010861. Epub 20230119. doi: 10.1371/journal.pcbi.1010861. PubMed PMID: 36656876; PubMed Central PMCID: PMCPMC9888713.
66. Shani-Narkiss H, Vinograd A, Landau ID, Tasaka G, Yayon N, Terletsky S, et al. Young adult-born neurons improve odor coding by mitral cells. Nat Commun. 2020;11(1):5867. Epub 20201117. doi: 10.1038/s41467-020-19472-8. PubMed PMID: 33203831; PubMed Central PMCID: PMCPMC7673122.
67. Gilday OD, Mizrahi A. Learning-Induced Odor Modulation of Neuronal Activity in Auditory Cortex. J Neurosci. 2023;43(8):1375-86. Epub 20230117. doi: 10.1523/JNEUROSCI.1398-22.2022. PubMed PMID: 36650061; PubMed Central PMCID: PMCPMC9987573.
68. Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, et al. Fully integrated silicon probes for high-density recording of neural activity. Nature. 2017;551(7679):232-6. doi: 10.1038/nature24636. PubMed PMID: 29120427; PubMed Central PMCID: PMCPMC5955206.
69. Pachitariu M, Steinmetz N, Kadir S, Carandini M, Kenneth D H. Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. bioRxiv. 2016:061481. doi: 10.1101/061481.
70. Shamash P, Carandini M, Harris K, Steinmetz N. A tool for analyzing electrode tracks from slice histology. BioRxiv. 2018:447995.
71. Vinje WE, Gallant JL. Sparse coding and decorrelation in primary visual cortex during natural vision. Science. 2000;287(5456):1273-6. doi: 10.1126/science.287.5456.1273. PubMed PMID: 10678835.
72. Green DM, Swets JA. Signal detection theory and psychophysics: Wiley New York; 1966.

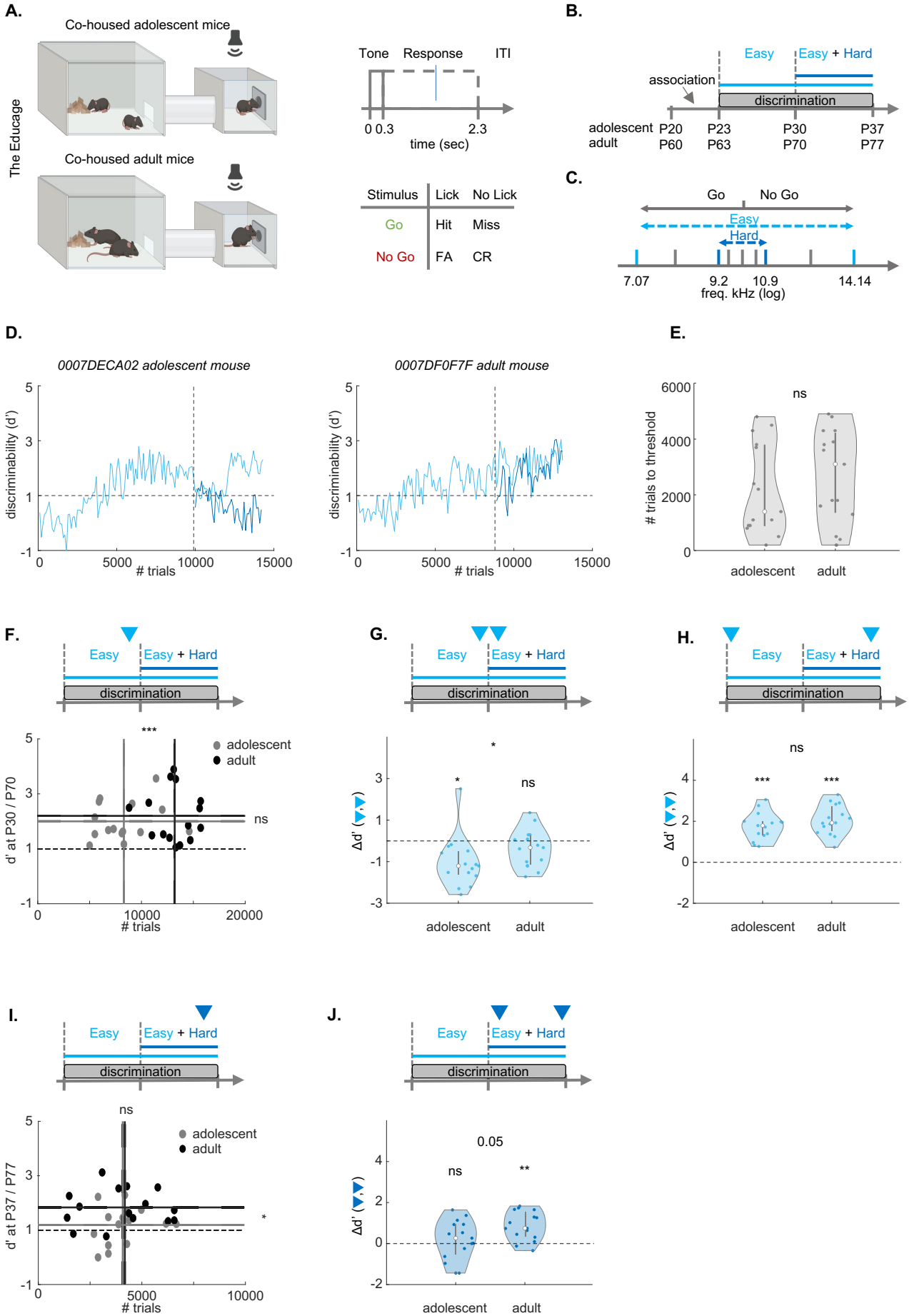

A.
Co-housed adolescent mice
The Educage
Co-housed adult mice
Tone
Response
ITI
0 0.3
2.3
time (sec)
Stimulus
Lick
No Lick
Go
Hit
Miss
No Go
FA
CR
B.
Easy
Easy + Hard
association
discrimination
adolescent
adult
P20
P23
P30
P37
P60
P63
P70
P77
C.
Go
No Go
Easy
Hard
7.07
9.2
10.9
14.14
freq. kHz (log)
D.
0007DECA02 adolescent mouse
0007DF0F7F adult mouse
discriminability (d')
trials
E.
trials to threshold
ns
adolescent
adult
F.
Easy
Easy + Hard
discrimination
d' at P30 / P70
adolescent
adult
***
ns
trials
G.
Δd' (▼,▼)
*
ns
H.
***
I.
d' at P37 / P77
*
J.
0.05
**

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). **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'>=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 ± 0.1791; # trials = 8317 ± 712) and adult mice at P70 (black; n=15, d'=2.1986 ± 0.2441; # trials = 13229 ± 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 (Δd') 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; Δd' 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; Δd' 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 ± 0.1783; # trials=4163 ± 297; adult-black; n=15, d'=1.8342 ± 0.1743; # trials=4102±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; Δd' between adult and adolescent mice: z = -1.9495; p = 0.0512, two-sample Wilcoxon rank sum test).

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') ~ age + sex + (1|cage).

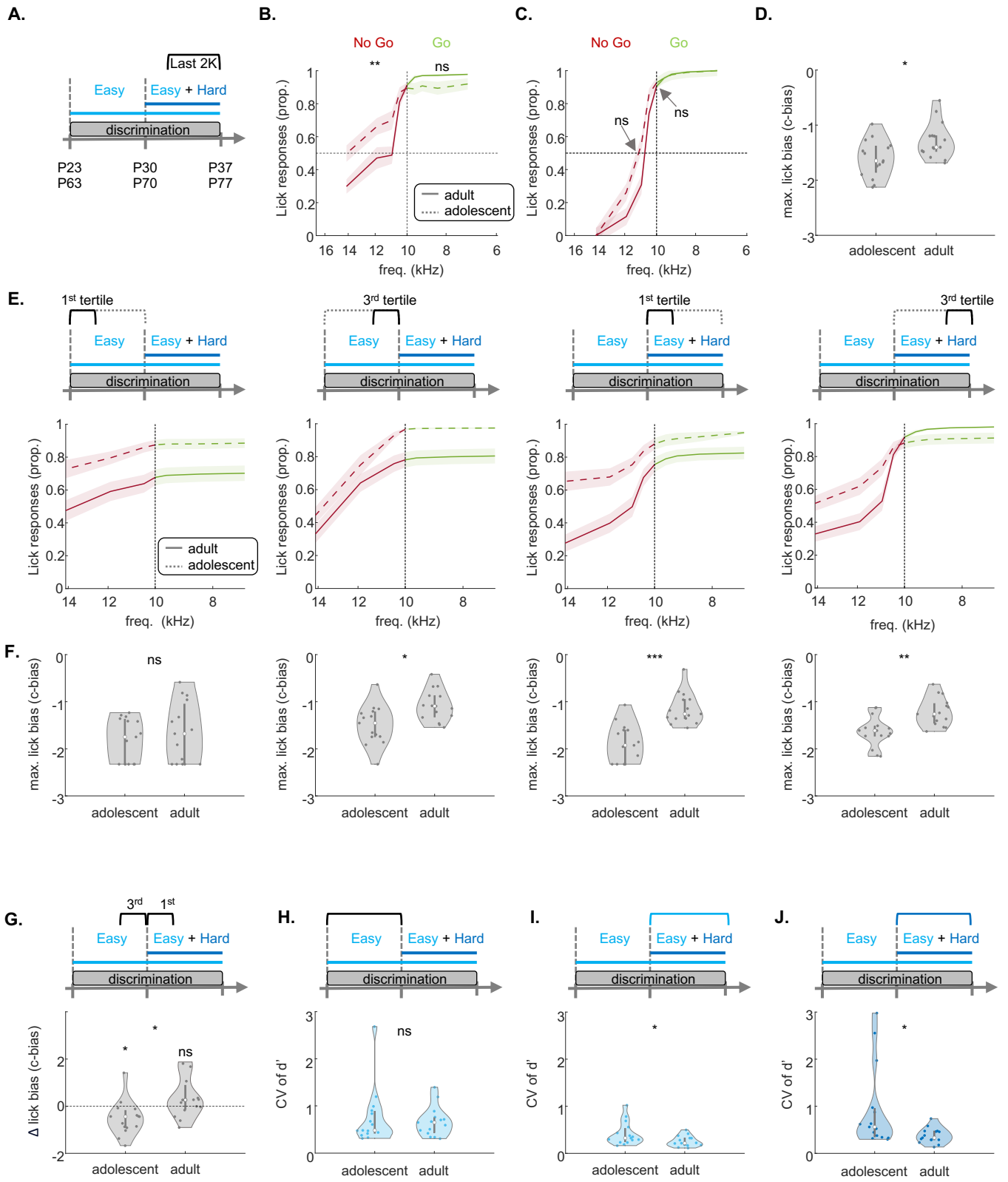

A.
Last 2K
Easy
Easy + Hard
discrimination
P23
P63
P30
P70
P37
P77
B.
No Go
Go
**
ns
Lick responses (prop.)
adult
adolescent
freq. (kHz)
C.
No Go
Go
ns
ns
Lick responses (prop.)
freq. (kHz)
D.
*
max. lick bias (c-bias)
adolescent
adult
E.
1st tertile
3rd tertile
1st tertile
3rd tertile
Easy
Easy + Hard
discrimination
Lick responses (prop.)
adult
adolescent
freq. (kHz)
F.
ns
*
***
**
max. lick bias (c-bias)
adolescent
adult
G.
3rd
1st
Easy
Easy + Hard
discrimination
*
*
ns
Δ lick bias (c-bias)
adolescent
adult
H.
ns
CV of d'
adolescent
adult
I.
*
CV of d'
adolescent
adult
J.
*
CV of d'
adolescent
adult

Figure 2*: Adolescent mice exhibit lower response inhibition and higher variability in performance over the course of learning.* **A.** Schematic timeline of tone discrimination. **B.** Average psychometric curve of the proportion of lick responses (above the threshold) across all stimuli (probe and learned; No-Go stimuli: $p = 0.0085$, Go stimuli: $p = 0.0963$, Wilcoxon rank sum test, after Bonferroni correction). **C.** Same as 'B', but normalized and fitted to a sigmoid function (proportion to the lick threshold at 10kHz, $p = 0.1163$; frequency at 50% lick rate, $p = 0.5896$, Wilcoxon rank sum-test, after Bonferroni correction). **D.** Maximal lick bias (i.e., criterion bias) per mouse during the last 2000 trials (c-bias: $p = 0.0136$, Wilcoxon rank sum test). **E.** Average psychometric curve, fitted to a sigmoid function across all stimuli (probe and learned) at different periods of the experiment — Left: first tertile-easy task (per mouse; adolescent $n = 15$, adult $n = 15$); Middle-left: last tertile – easy task; Middle-right: first tertile-hard task; Right: last tertile-hard task. (Dashed vertical =category boundary). **F.** Same as E. as the maximal C-bias. From left to right:($z = -0.5104$, $p = 0.6098$, two-sample Wilcoxon rank sum test; $z = -2.6546$, $p = 0.0079$, two-sample Wilcoxon rank sum test; $z = -4.0739$, $p = 4.6228e-05$, two-sample Wilcoxon rank sum test; $z = -3.2768$, $p = 0.0011$, two-sample Wilcoxon rank sum test). **G.** Change in lick bias (Δ c-bias) between the last 100 trials before and after the introduction of the hard task. (left, adolescent signed rank = 17, $p = 0.0125$; right, adult signed rank = 83, $p = 0.2078$, one-sample Wilcoxon signed rank test; Δ c-bias between adolescents and adults, $z = -2.9035$; $p = 0.0037$, two-sample Wilcoxon rank sum test). **H.** Coefficient of variation (CV) of the d' of the easy task before the introduction of the hard task. ($z = -0.1659$, $p = 0.8682$, two-sample Wilcoxon rank sum test) **I.** Same as 'H' after the introduction of the hard task ($z = 2.1569$, $p = 0.0310$, two-sample Wilcoxon rank sum test). **J.** Same as 'H' but for d' of the hard task ($z = 2.5302$, $p = 0.0114$, two-sample Wilcoxon rank sum test).

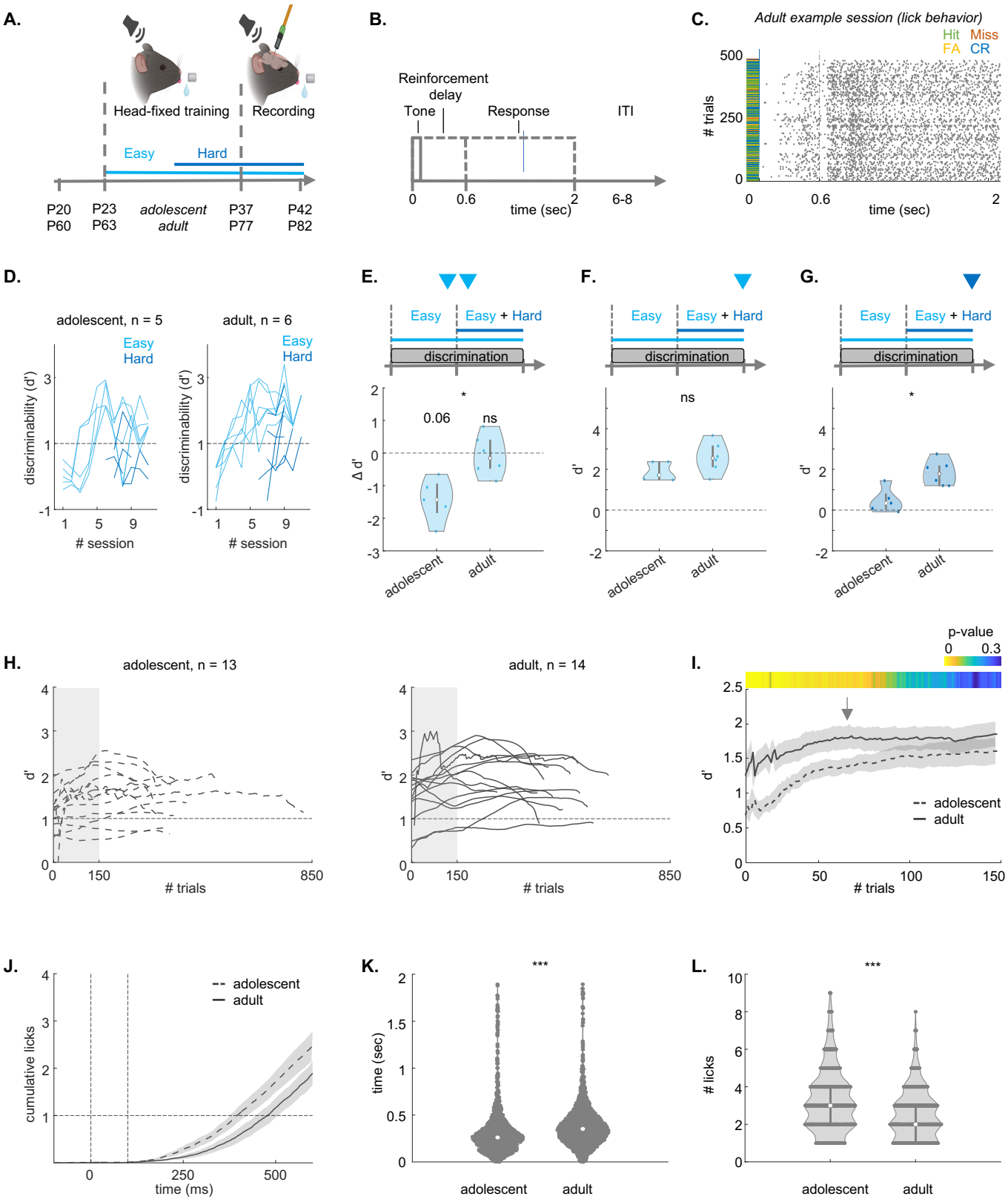

A.
Head-fixed training
Recording
Easy
Hard
P20
P60
P23
P63
adolescent
adult
P37
P77
P42
P82
B.
Reinforcement delay
Tone
Response
ITI
0
0.6
2
6-8
time (sec)
C.
Adult example session (lick behavior)
Hit
Miss
FA
CR
500
250
0
trials
0
0.6
2
time (sec)
D.
adolescent, n = 5
adult, n = 6
Easy
Hard
discriminability (d')
session
E.
F.
G.
Easy
Easy + Hard
discrimination
0.06
ns
*
Δ d'
d'
adolescent
adult
H.
adolescent, n = 13
adult, n = 14
d'
trials
I.
p-value
0
0.3
d'
adolescent
adult
trials
J.
adolescent
adult
cumulative licks
time (ms)
K.
***
time (sec)
adolescent
adult
L.
***
licks
adolescent
adult

Figure 3: *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.

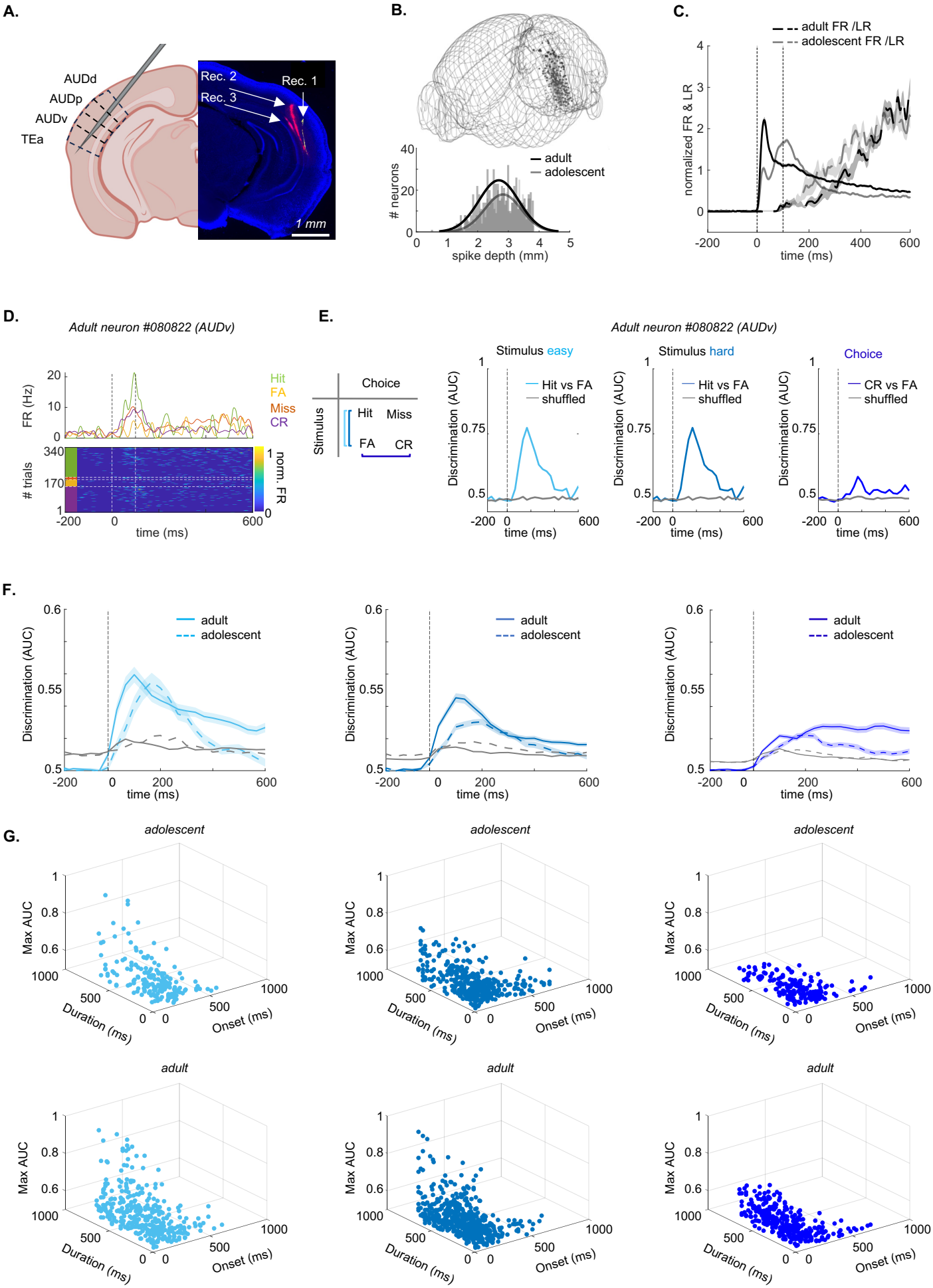

A.
AUDd
AUDp
AUDv
TEa
Rec. 2
Rec. 1
Rec. 3
1 mm
B.
neurons
adult
adolescent
spike depth (mm)
C.
adult FR /LR
adolescent FR /LR
normalized FR & LR
time (ms)
D.
Adult neuron #080822 (AUDv)
FR (Hz)
Hit
FA
Miss
CR
trials
norm. FR
time (ms)
E.
Adult neuron #080822 (AUDv)
Choice
Stimulus
Hit
Miss
FA
CR
Stimulus easy
Stimulus hard
Choice
Hit vs FA
shuffled
CR vs FA
Discrimination (AUC)
time (ms)
F.
adult
adolescent
Discrimination (AUC)
time (ms)
G.
adolescent
adult
Max AUC
Duration (ms)
Onset (ms)

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). **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 -200ms to +600ms 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 -200ms to 600ms 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 (+- 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).

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|Rec ID); duration (ms) ~ age + (1|Mouse ID) + (1|Rec ID); maximal discriminability (AUC) ~ age + (1|Mouse ID) + (1|Rec ID).

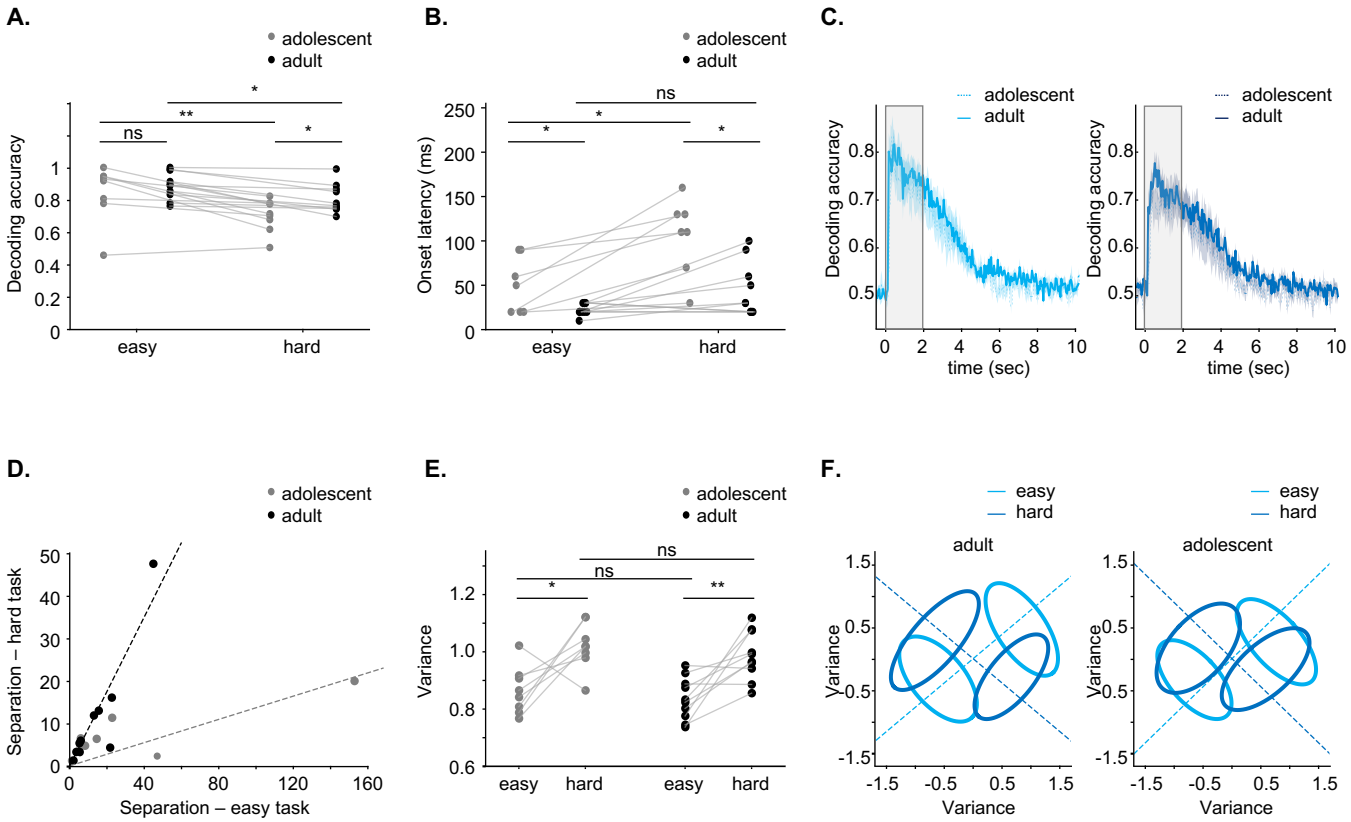


Figure 5: *Decoding in adult neuronal populations outperforms decoding in adolescents.* **A.** Decoding accuracy for the first 200ms across all recordings in both adults (black) and adolescents (grey) for the easy task (adolescents compared to adults, p = 0.5000, Student's t-test) and the hard task (adolescents 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, pared t-test; hard task: p = 0.0100). **C.** Decoding accuracy over a time window from -0.5s to 10s (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.

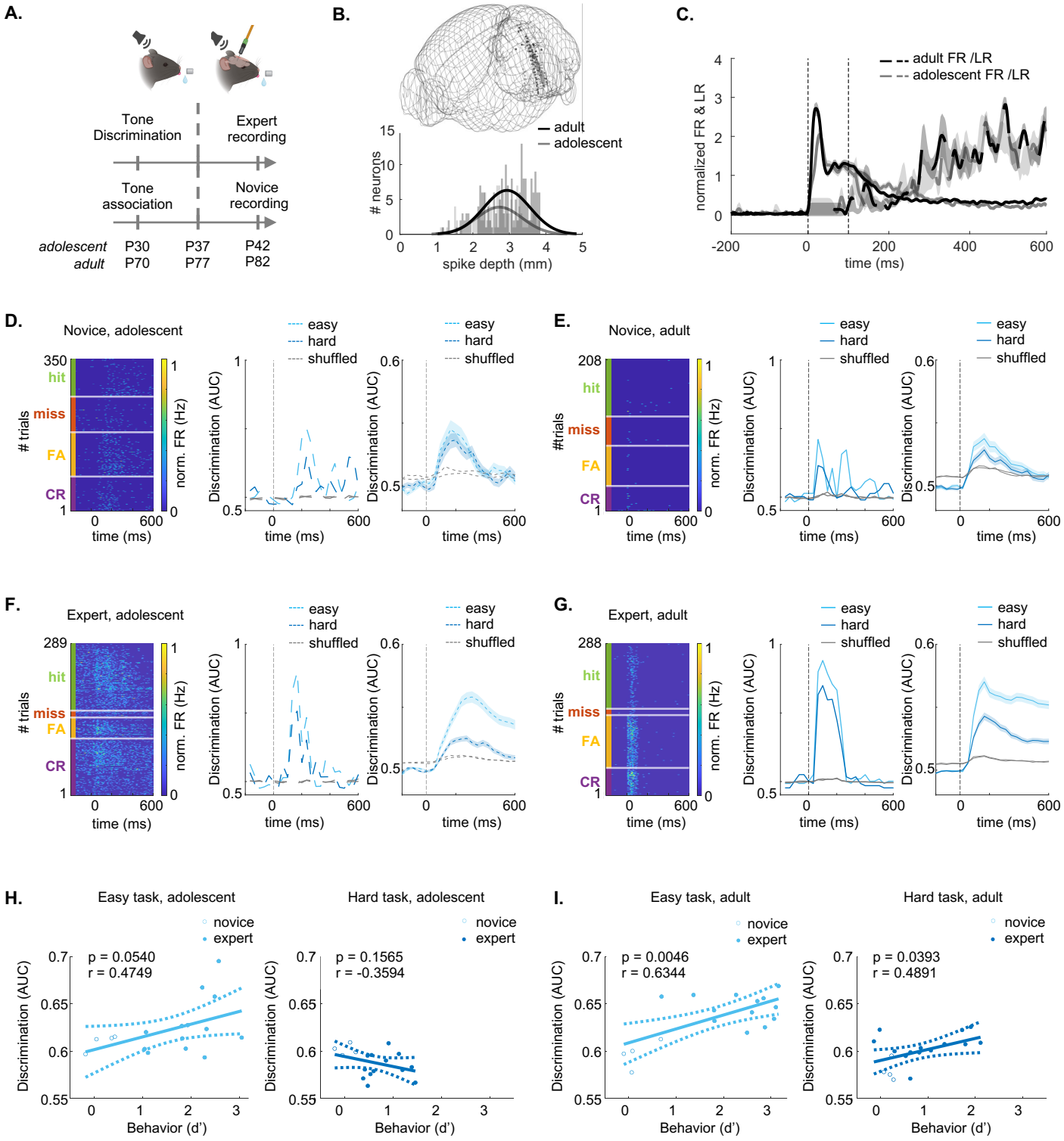


Figure 6: *Cortical activity during behavior reflects both age- and learning-induced effects.* **A.** Training and recording schedule for novice mice, compared to expert mice. **B.** 3D-Reconstruction of recording sites in novice adolescent (n = 6; grey) and novice adult (n = 6; black) mice. Bottom: 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 +-STE. **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 (+-STE) 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 includes 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 10). Model structures: onset latency (ms) ~ age* learning * difficulty + (1|Mouse ID) + (1|Rec ID); duration (ms) ~ age* learning * difficulty + (1|Mouse ID) + (1|Rec ID); maximal discriminability (AUC) ~ age* learning * difficulty + (1|Mouse ID) + (1|Rec ID).

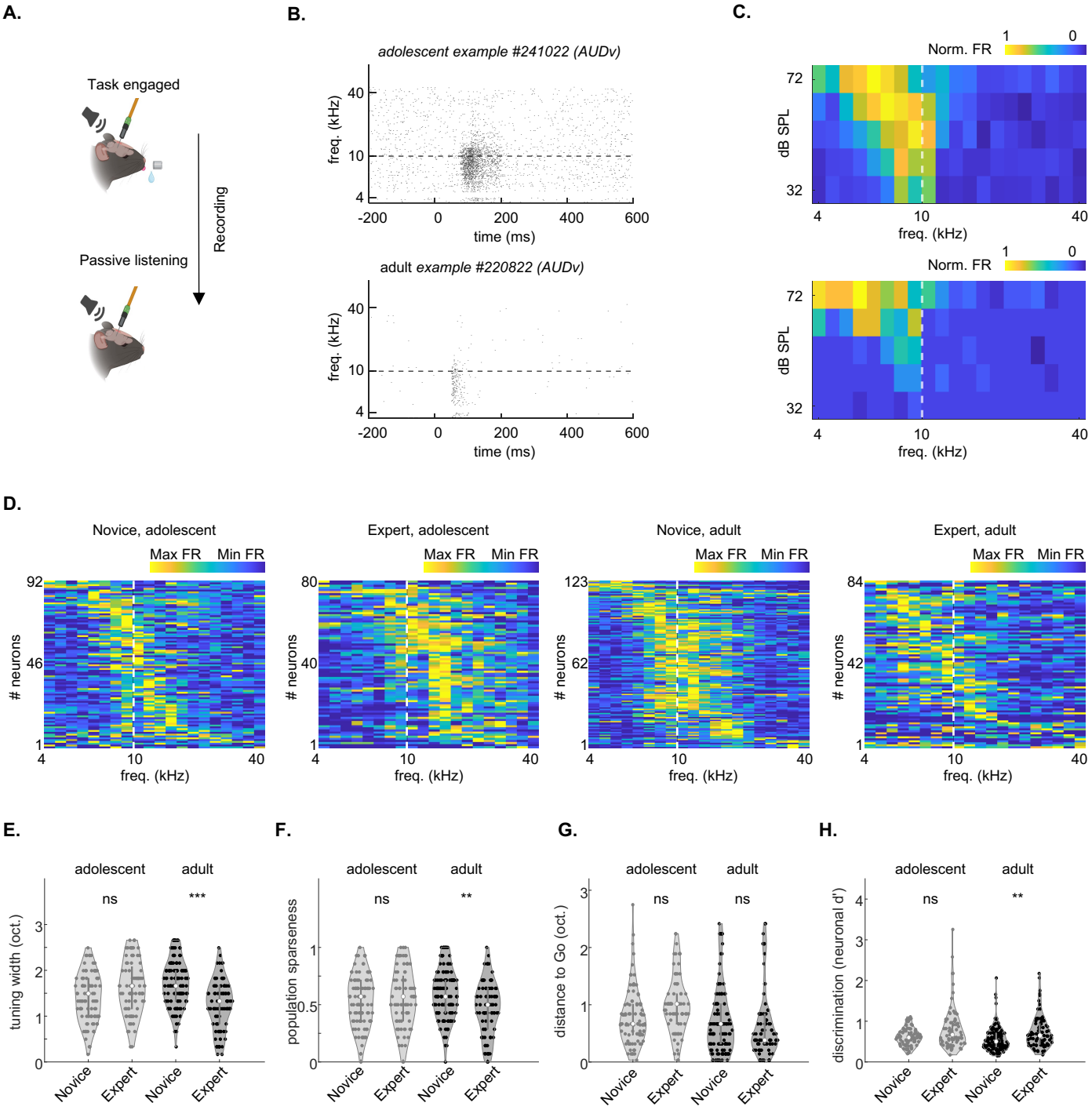
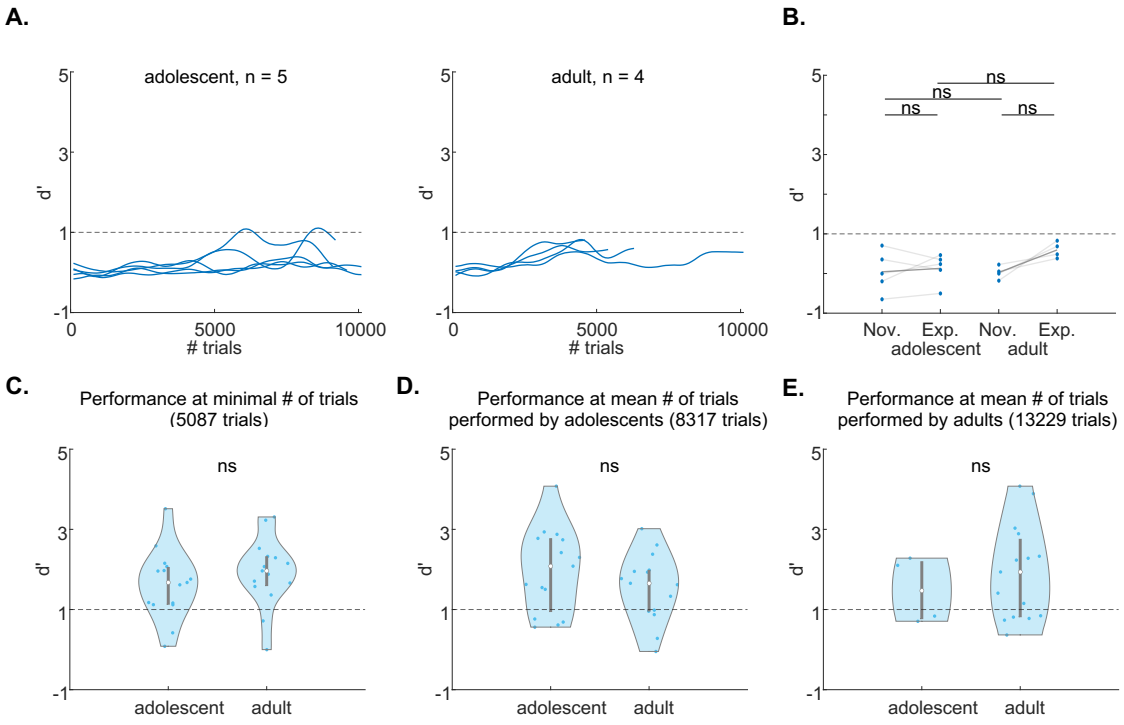
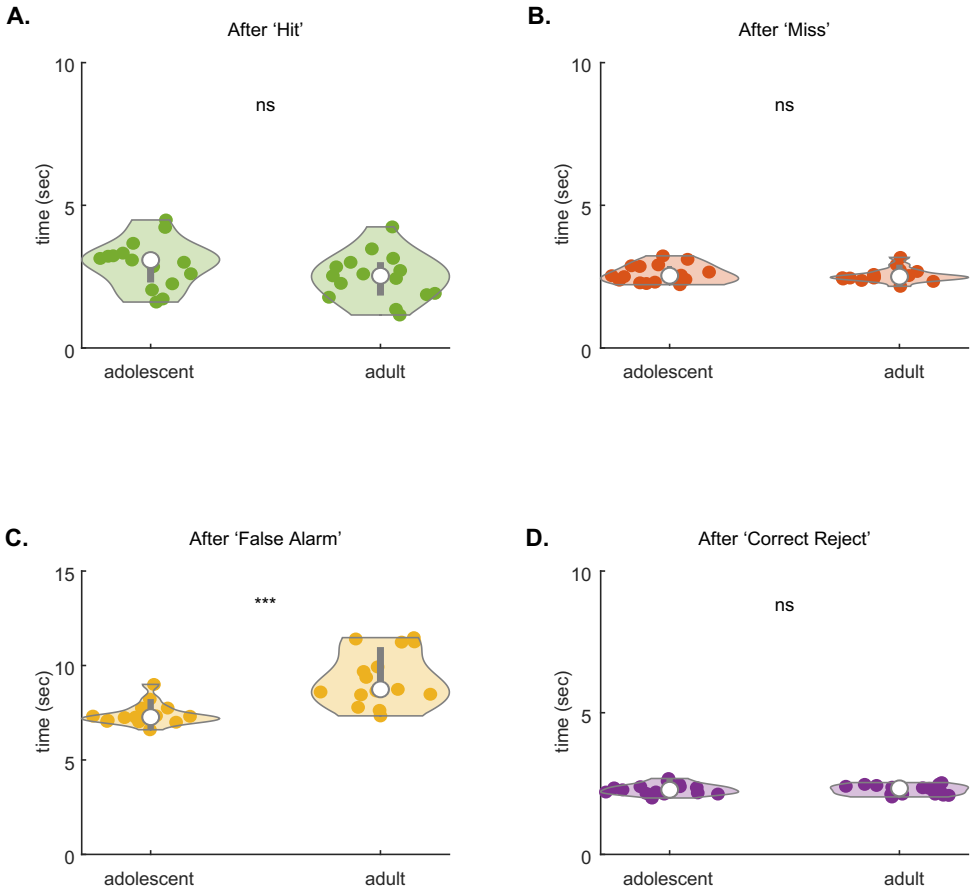


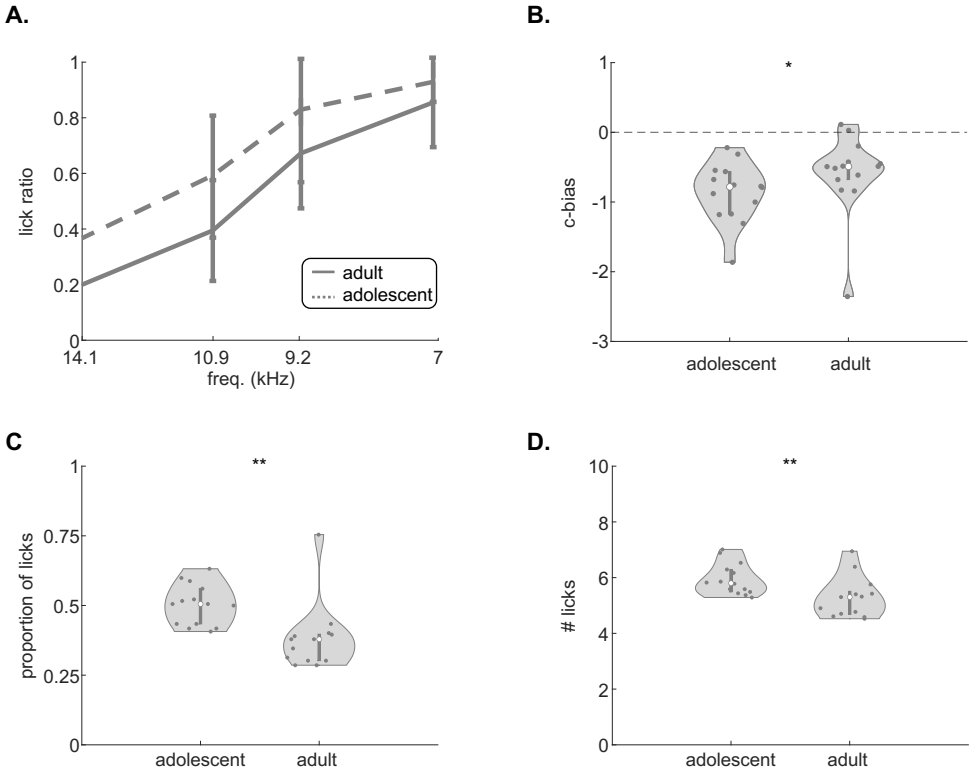
Figure 7: *Adult mice show greater plasticity after learning.* ***A.*** Schematic showing that for the passive listening protocol, we continued our recording following the session of the engaged task (i.e. in satiated mice) by removing the waterspout. **B.** Example raster plot of a neuron from an adolescent 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 the 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).



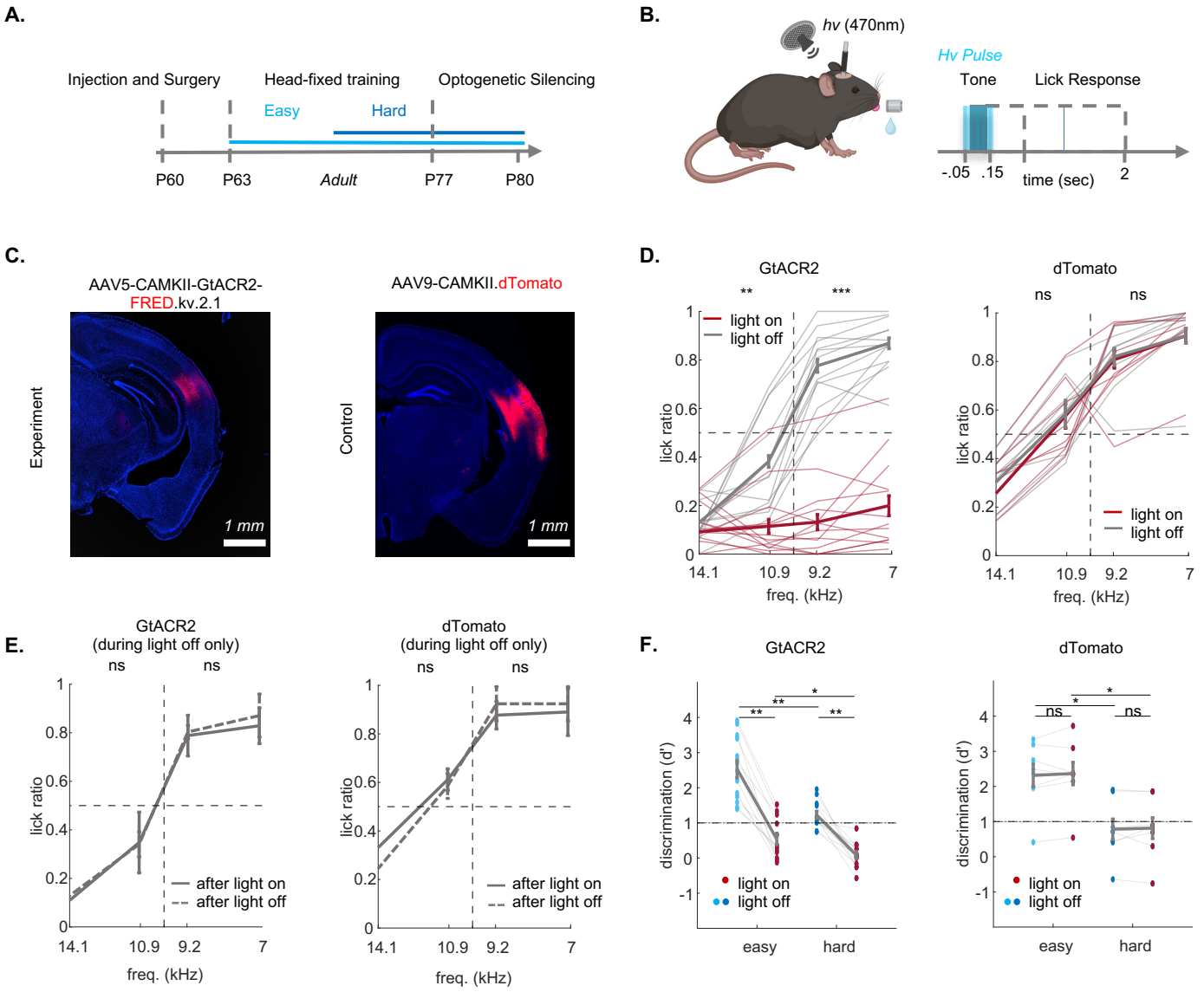
Supplementary Figure 1-1: *Behavioral criteria of auditory learning.* **A.** Learning curves per mouse throughout the experiment (n=5 adolescents, n=4 adults). **B.** Discriminability of novice (Nov.; first 100 trials) compared to expert (Exp.; last 100 trials) mice in the adolescent and adult groups (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).



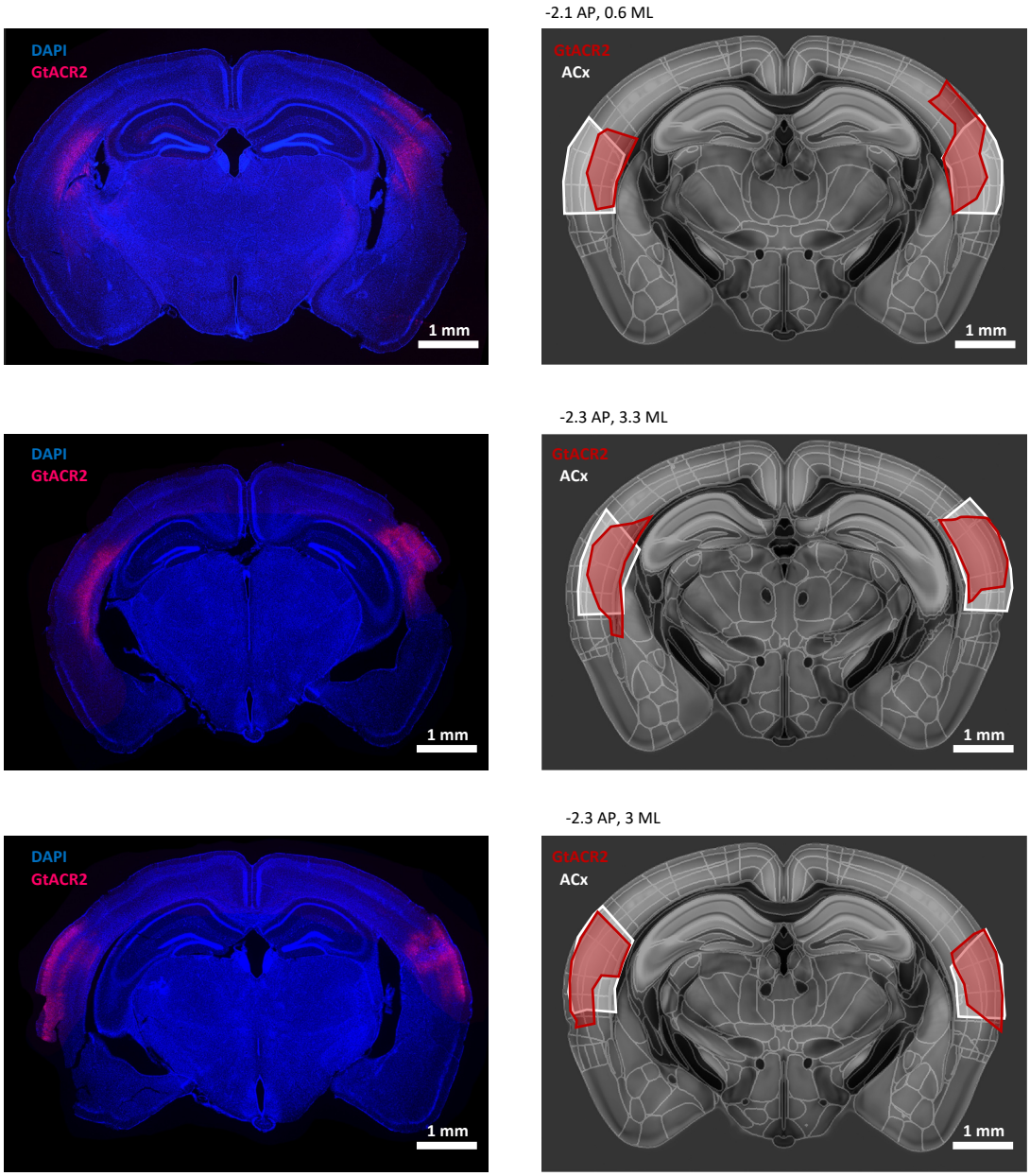
Supplementary Figure 2-1: *Inter trial interval after different trial outcomes.* **A.** Average *i*nter trial interval (ITI) per mouse to the next trial after a previous hit (z = 1.6176; p = 0.1057, two-sample Wilcoxon rank sum test). **B.** Same as A. after a previous miss (z = 0.0830; p = 0.9339, two-sample Wilcoxon rank sum test). **C.** Same as A. after a previous false alarm hit (z = -3.9823; p = 6.8241e-05, 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).



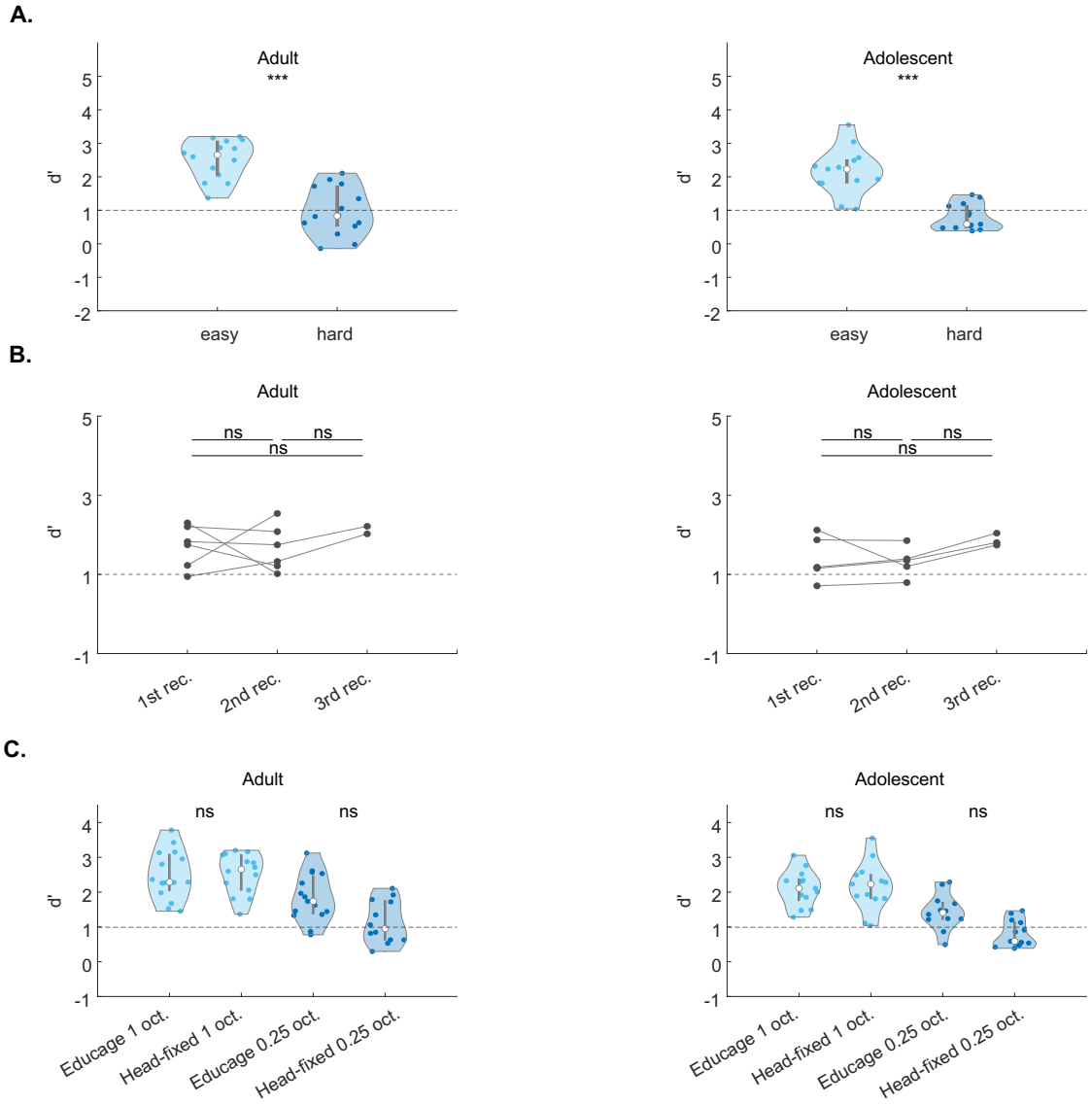
Supplementary Figure 3-1: *Lick bias and impulsivity in adolescent and adult mice during head-fixed recordings.* **A.** Average psychometric curve of adult recordings (solid line) and adolescent recordings (dashed line) **B.** Lick bias (i.e., criterion bias) per recording (c-bias: z = -2.1366, p = 0.0326, 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). **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).



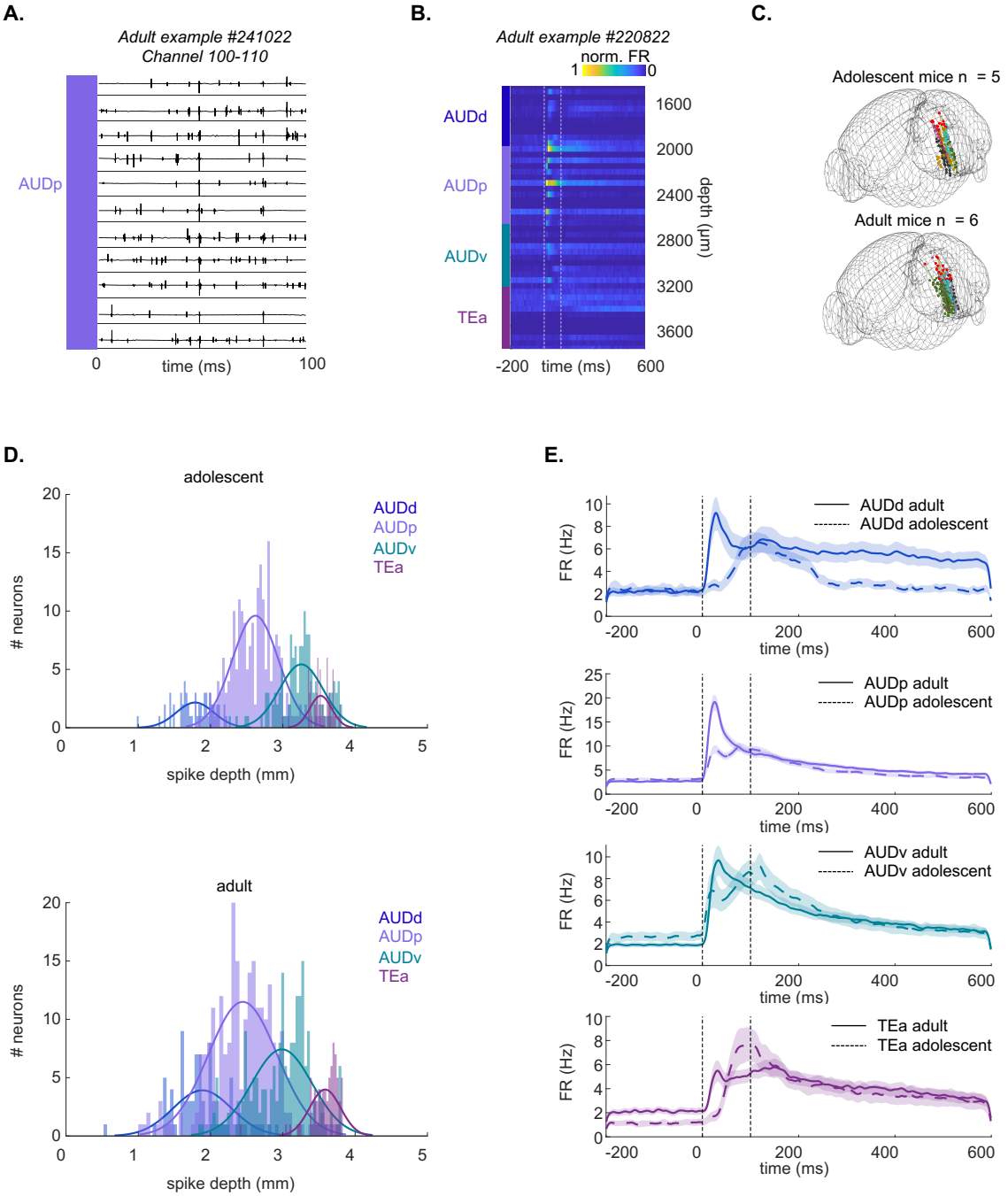
Supplementary Figure 3-3: *Verification of GtACR2 expression*. Photomicrograph (left), and reconstruction (right) of the GtACR2 infected area (red) per mouse. Reconstruction followed the coordinates of the Allen-CFF template-atlas. ACx is highlighted in white and GtACR2-expression in red.



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 = 2.4414e-04; adult mice: signed rank: 105, p = 1.2207e-04, 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.

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 * dprime + (1|Mouse ID) + (1|Recording ID).



Supplementary Figure 4-1: *Probe reconstruction and activity profile across ACx regions.* **A.** Example voltage trace during tone (100ms) across 11 channels. **B.** Example PSTH per depth (50 µm bins) across AUDd, AUDp, AUDv and TEa (3850 maximal depth µm). ***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).

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).

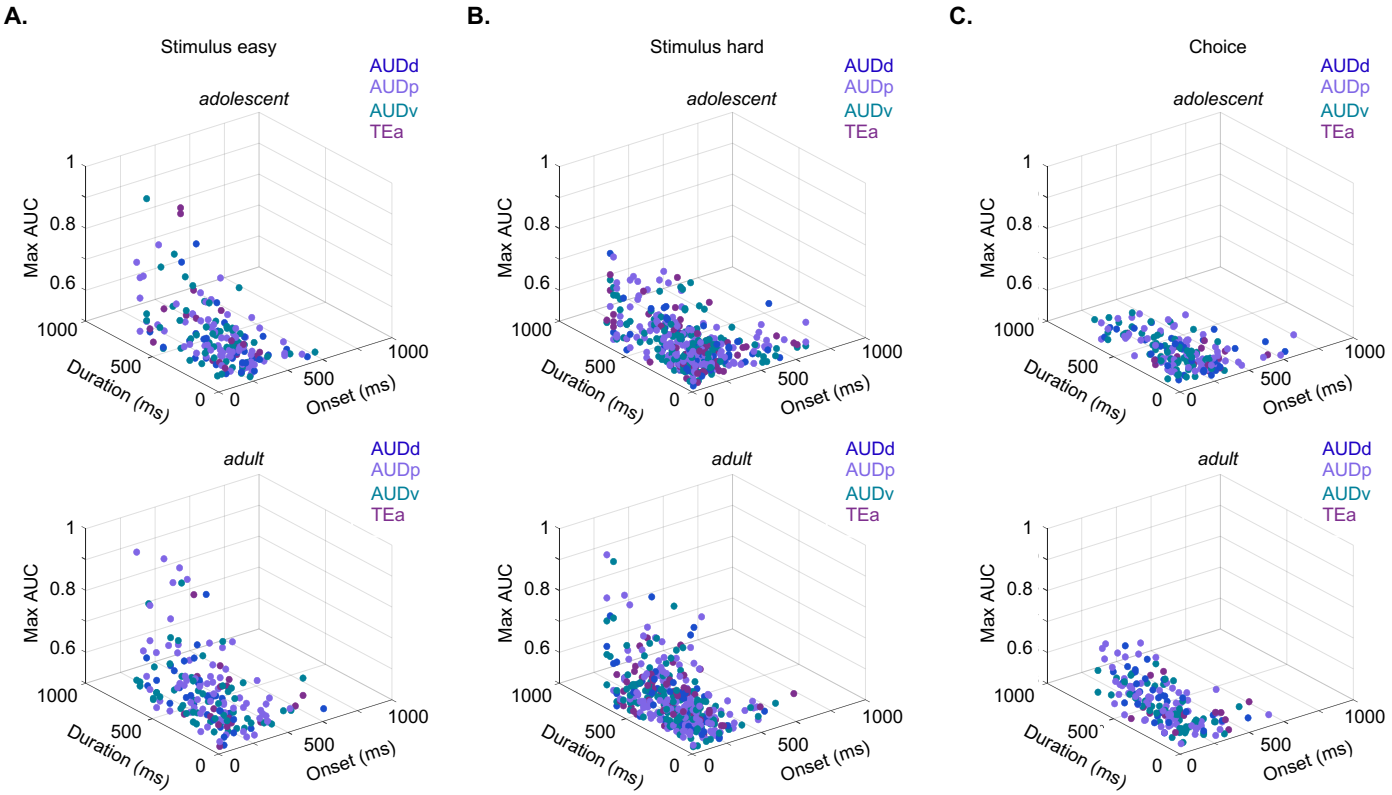
AUDd	adolescent		adult					
neuronal property	Mean	± STE	Mean	± STE	Effect size	lower CI	upper CI	p-value
spontaneous FR	2.4274	0.4188	2.2271	0.3668	0.1414	-0.0111	0.2997	0.9801
evoked FR	14.5588	1.6046	22.1731	2.9083	-0.306	-0.3798	-0.1396	0.1537
FR coeff. var	2.4484	0.2496	1.7209	0.1238	0.4794	0.0573	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 (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).

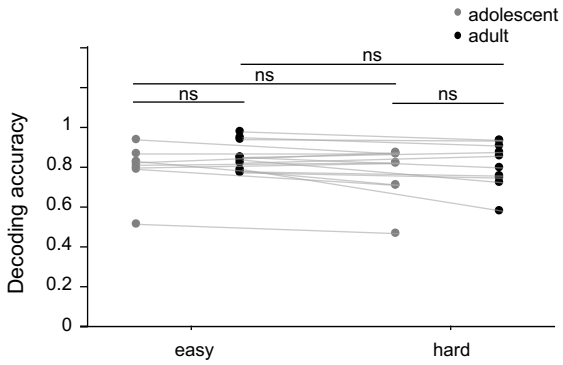
adolescent	P-Value**					
neuronal property	AUDd - AUDp	AUDd - AUDv	AUDd - TEa	AUDp - AUDv	AUDp - TEa	AUDv - TEa
spontaneous FR	0.3379	0.9998	0.3018	**0.0793**	**0.0004**	0.189
evoked FR	**0.0276**	0.5874	0.9983	0.2043	**0.0367**	0.6856
FR coeff. var	**0.0078**	0.3519	0.5517	0.2063	0.3573	0.9992
latency 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

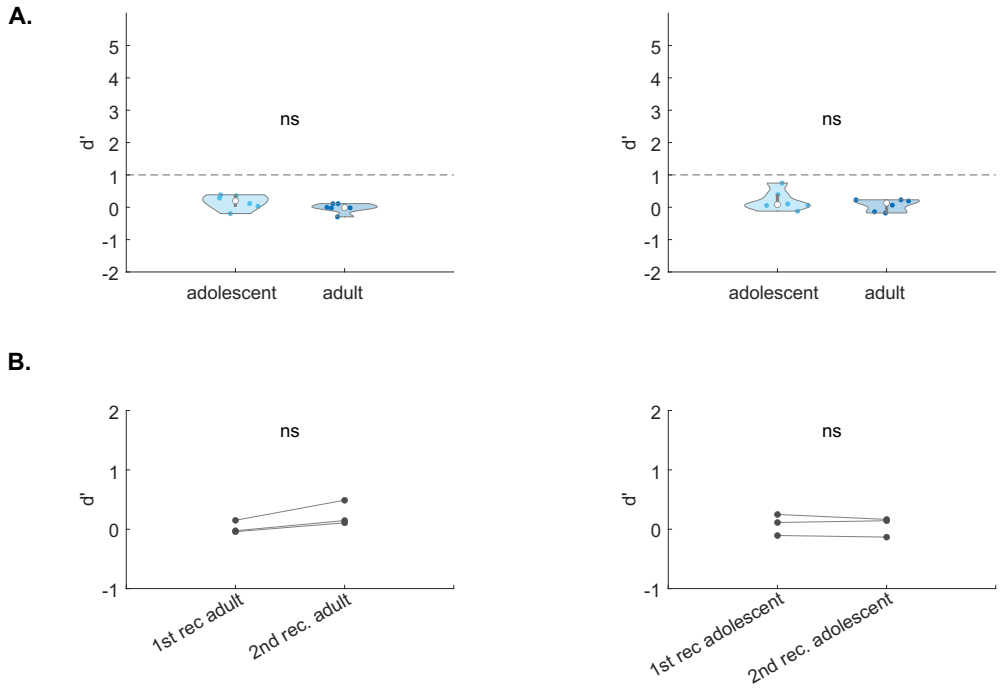
Supplementary Table 4: *Adolescent and adult firing properties of expert mice are distinct between different sub-regions*. P-values of Kruskal Willis Test after Tukey-Kramer correction for multiple comparisons of the average baseline FR (Hz), evoked FR (Hz), coefficient of variance of FR, latency to 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).



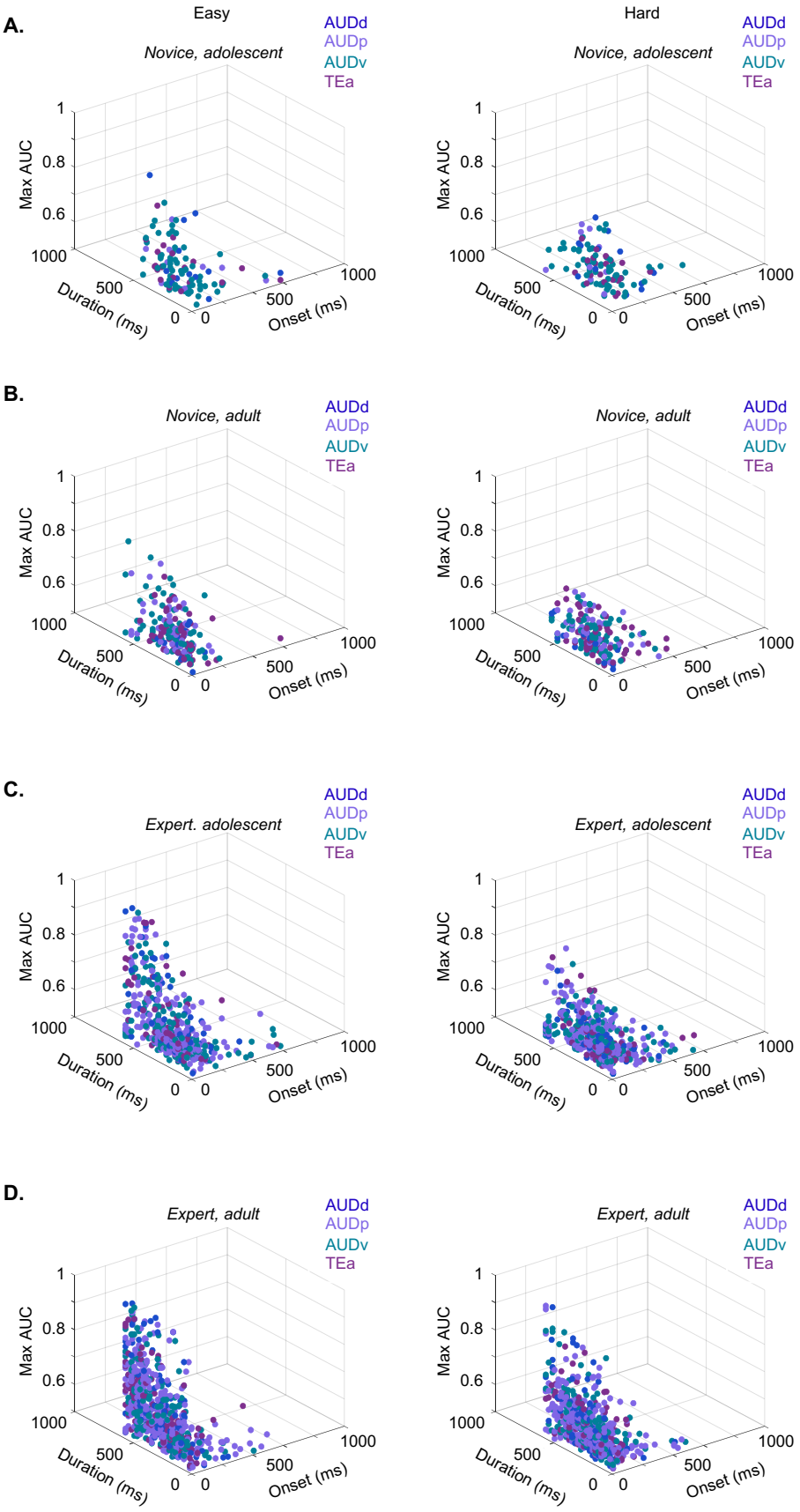
Supplementary Figure 4-2. *The neuronal discriminability of stimulus- and choice related activity is similar across auditory sub-regions.* **A.** Onset-latency of discriminability (ms), duration of discriminability (ms), and maximal discriminability (AUC) of neurons that showed significant discriminability (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, AUDv, 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.


A.
Easy
Novice, adolescent
Hard
Novice, adolescent
B.
Novice, adult
Novice, adult
C.
Expert. adolescent
Expert, adolescent
D.
Expert, adult
Expert, adult
AUDd
AUDp
AUDv
TEa
Max AUC
1
0.8
0.6
Duration (ms)
1000
500
0
Onset (ms)
0
500
1000

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* **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 novice, adolescent mice across the AUDd, AUDp, AUDv, TEa. Left: easy task (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; Friedman test, correct for multiple comparisons). Right: hard task (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; Friedman test, correct for multiple comparisons). **B.** Same as 'A' for novice, adult mice. Left: easy task (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; Friedman test, correct for multiple comparisons). Right: hard task (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; Friedman test, correct for multiple comparisons). **C.** Same as 'A' for expert, adolescent mice. Left: easy task (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; Friedman test, correct for multiple comparisons). Right: hard task (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; Friedman test, correct for multiple comparisons). **D.** Same as 'A' for expert, adult mice. Left: easy task (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 (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).

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

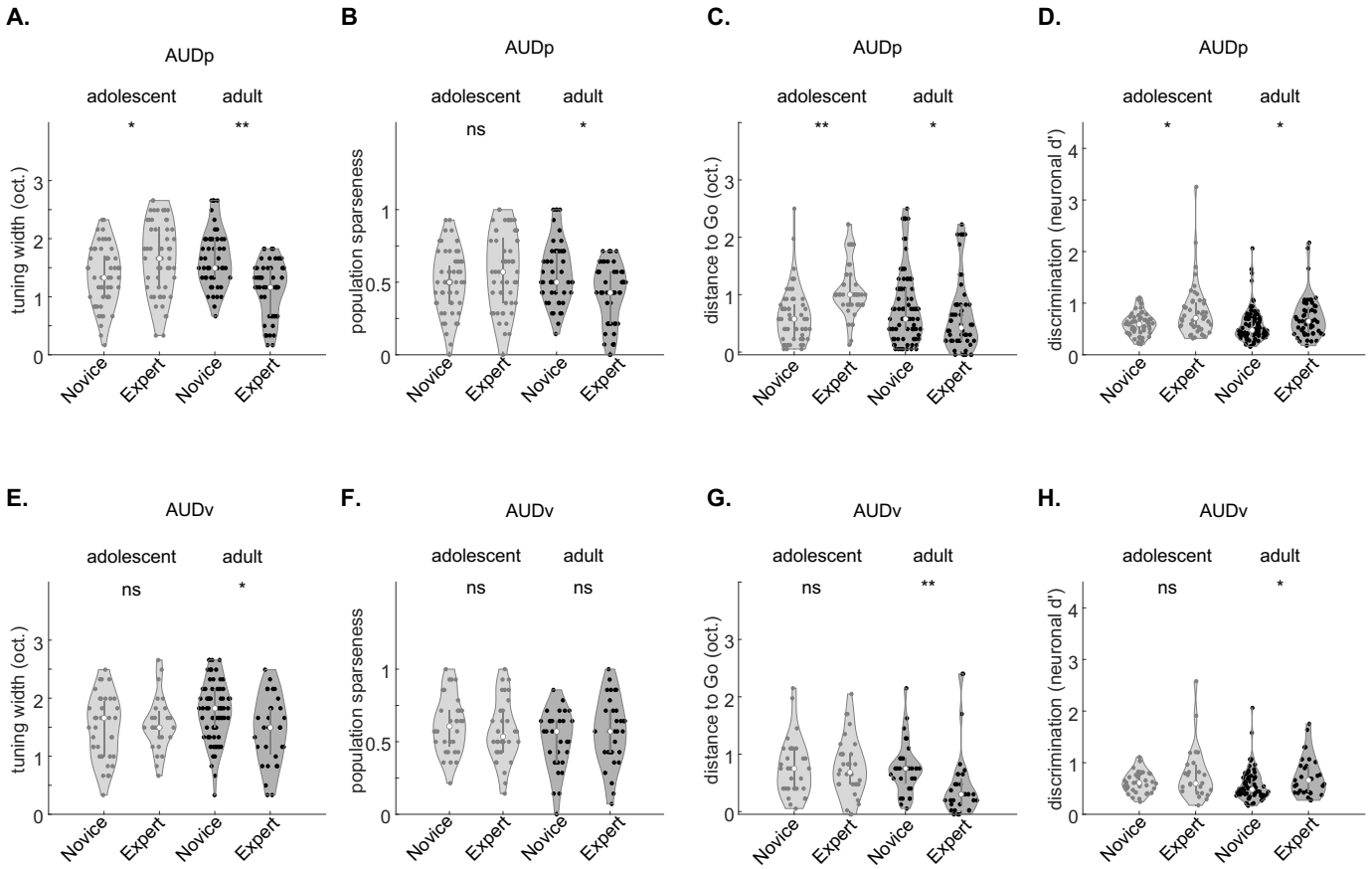
Supplementary Table 5: *Novice, adolescent and adult mice have distinct firing properties in different sub-regions***.** Mean and standard error (STE), mean effect size (robust Cohen's D), lower and upper Confidence Interval (CI) and p-value (Wilcoxon rank-sum test, adjusted for multiple comparisons 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).



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.

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.