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 20-40) and adult (postnatal day 60-80) 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 the harder task. This performance difference was due to higher response variability and weaker cognitive control expressed as lower response 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 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 higher 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, 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 distinct in 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 (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 first 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), there was no significant difference between the performance of adult and adolescent mice at the end of the first week (**Fig. 1f**, horizontal lines; d' calculated from the last 100 trials). Following one week of training on the easy task, we introduced the second pair of pure tones (**Fig. 1b**; 'Easy +

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

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

124 Although all mice trained for exactly 14 days, each mouse performed a unique number
125 of trials, as these were initiated by the mice spontaneously. As noted above, pre-adolescent
126 mice performed fewer trials/day as compared to adults but, notably, only during the first week
127 of training (mean ± SEM: adolescents, 8317 ± 712 trials, adults 13229 ± 514 trials, t-stat = -
128 5.6314, df = 28, p = 4.9566e-06, two-sample independent t-test). Still, this difference raises the
129 question whether the different number of trials affected our conclusions from the first week of
130 training. Thus, to evaluate the possible connection between the number of trials and
131 performance in our data, we carried out several analyses. First, we found that the number of
132 trials and d' were not correlated in either age-group during the first week of discrimination on
133 the easy task (Pearson-r; adolescent: r = 0.1880, p = 0.5023; adult: r = -0.0730, p = 0.7959), nor
134 on the second week of discrimination on both the easy and the hard task (Pearson-r; adolescent:
135 r = 0. 2264, p = 0. 4171; adult: r = 0. 0194, p = 0. 9454). Second, we compared performance at
136 three time points along the task with reference to times with shared number of trials: 1)
137 evaluating d' after the minimal number of trials performed by all mice during the first week of
138 training, and 2) evaluating d' at the mean number of trials of each group. We found no
139 significant differences between adolescents and adults in all comparisons (**Fig. S1-1c-e**).

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

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

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

187

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

189 To enable electrophysiological recordings during expert task performance, we trained
190 adolescent (n = 5) and adult (n = 6) mice on a head-fixed learning paradigm, using a similar
191 protocol as in the Educage (**Fig. 3a**). In the head-fixed protocol, water access was limited to
192 training sessions. Head-fixed mice were trained to lick after a 100ms tone and were rewarded
193 (with water) or punished (by a 2 sec white noise) after a 600ms delay (**Fig. 3b**). Each session
194 was concluded after the mice became satiated and stopped licking. Water supply was limited to
195 0.0125 ml per day per gram of body weight, and mice that did not consume this amount were

196 compensated after training (see Methods) (**Fig. 3c** shows an example session from one mouse).
197 The number of trials per training session was significantly lower in adolescents (mean ± SEM:
198 adolescents: 410 ± 35, adults: 580 ± 37 t-statistic = -27.9855, p = 0.0012, independent two-
199 sample t-test). However, the number of sessions to reach the behavioral threshold (d' >= 1 in
200 the easy task) was not significantly different between adolescents and adults (mean ± SEM:
201 adolescents: 3.7 ± 0.8, adults: 3.8 ± 0.7; t-statistic = - 0.8531, p = 0.1971, independent two-
202 sample t-test, **Fig. 3d**). While the number of trials performed within each session is likely
203 related to physiological factors, such as smaller body weight and faster satiation, they did not
204 impact learning capabilities. Similar to learning in the Educage, adolescent performance (but
205 not that of adults) decreased after we introduced the hard task (**Fig. 3e**). Another similarity
206 between the head-fixed and Educage versions of the task was that we found no significant
207 differences between adolescents and adults in the performance of the easy task (**Fig. 3f**), and
208 that adults outperformed adolescents on the hard sound discrimination (**Fig. 3g**). Also similar
209 to the behavior in the Educage, lick rates of adolescent mice were generally higher (**Fig. S3-**
210 **1a**), and adolescents exhibited a stronger lick bias (**Fig. S3-1b**). In contrast, in the head-fixed
211 task we found no evidence in impulsiveness; there were no differences between the groups in
212 the licking patterns during the inter-trial intervals (**Fig. S3-1c, d**).

213 To test whether ACx was necessary for expert discrimination in this task, we performed
214 a causal experiment in adult mice. Adult mice (n=3) were injected bilaterally with AAV5-
215 CAMKII-GtACR2-FRED_kv_2.1 into the ACx, and optical fibers were implanted over the
216 injection sites (**Fig. S3-2a,c**; **Fig. S3-3**). This viral construct enables optogenetic silencing of
217 excitatory neurons via GtACR2 (*Guillardia theta* Anion Channelrhodopsin 2), under the
218 control of the CaMKII promoter. It includes a RED Fluorescent protein for expression
219 verification (FRED) and a soma-targeting motif derived from the Kv2.1 potassium channel to
220 enhance membrane localization. Mice were trained identically on the head-fixed Go/No-Go
221 paradigm as described above. At the end of training, we attached a light source to the implanted
222 fibers to allow optogenetic suppression during several sessions (**Fig. S3-2a**). In these sessions,

223 light was applied in 50% of trials in a pseudo-random fashion, with light ON trials starting from
224 -50ms prior to tone onset up to 50ms after tone offset (**Fig. S3-2b**). A control group of mice
225 (n= 3) were injected with AAV9-CaMKII-dTomato and went through the exact same procedure
226 (**Fig. S3-2a,c**). We compared the lick rates for both Go and No-Go stimuli under light ON and
227 light OFF conditions. We found that lick rates were strongly affected in the GtACR2-injected
228 mice (n = 13 sessions), but not in controls (n = 8 sessions) (**Fig. S3-2d**). Light ON conditions
229 only affected the lick rate of the inhibited trial and not the follow-up trials (**Fig. S3-2e**).
230 Discriminability (d') decreased significantly with optogenetic suppression for both easy and
231 hard tasks, although mice still performed better in the easy task under both conditions (**Fig. S3-**
232 **2f**). These results confirm findings by others in similar (though not identical) tasks [19, 20],
233 suggesting that the ACx is necessary for executing the behavior during the expert stage of this
234 task (but see Discussion).

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

247 We found additional differences between the behavior of adults and adolescents in the
248 head restrained configuration. The average adolescent discriminability was lower at the
249 beginning of the session (initial 78 trials) but then rapidly leveled out (**Fig. 3i**). Adolescent mice
250 had shorter lick latencies (**Fig. 3j, k**), and higher lick counts (**Fig. 3j, l; Supplementary Table**

251 **1**). While the lick latencies were independent of the discriminability (d'), we found a significant
252 interaction effect between lick count and the d' (**Supplementary Table 1**). Together, these
253 behavioral differences suggest that the weaker response inhibition (here, expressed as lick
254 latency) and higher reward anticipation (here, expressed as lick count) may contribute to the
255 weaker performance of adolescents during the beginning of the recording.
256
257 *Neuronal representations of stimulus- and choice-related activity are immature in adolescents*
258 To compare the activity of neurons in adolescent and adult mice, we recorded spiking activity
259 from the ACx using the high-density Neuropixels-1 probes in expert mice engaged in the task.
260 We targeted the ACx by inserting the probe in a diagonal angle, traversing four auditory
261 regions: Dorsal Auditory Cortex (AUDd), Primary Auditory Cortex (AUDp), ventral auditory
262 cortex (AUDv), and Temporal Association Cortex (TEa) (**Fig. 4a** left**, Fig. S4-1a-b** show an
263 example voltage trace and one example recording). We recorded multiple times from each
264 mouse and used DiI- or DiO-coatings for different penetrations to verify probe trajectories
265 postmortem. We used the 3D-allen CCF-slice reconstruction (github.com/cortex-
266 lab/allenCCF), for high resolution anatomical registration (**Fig.4a, b**; **Fig. S4-1c, d**).
267 We limited our single neuron analysis to well-isolated single units from infragranular
268 layer 5 and layer 6 (recordings were restricted to these layers due to the probe angle), as two
269 main projection layers of the cortex, which are key nodes for assessing cortical outputs [21].
270 Recorded neurons from other layers and areas were excluded from the analysis. Together, we
271 collected data from 1145 single neurons in 13 recordings of 5 adolescent mice and 1267 single
272 neurons in 14 recordings of 6 adult mice (**Supplementary Table 2;** an overview of every
273 dataset can be found in **Supplementary Table 7**). We further restricted our analysis to single
274 neurons excited by sounds during the first 150ms after tone onset, to exclude any contributions
275 of motor actions due to licking (**Fig. 4c**; **Fig. S4-1e**). Lick rate- and Firing rate-PSTHs were not
276 correlated during the first 150ms (adolescents $r = -0.1515$, $p = 0.9878$; adults $r = -0.0638$, $p =$
277 0.9942). In total, we analyzed $n = 463$ neurons from adolescent mice and $n = 599$ from adults.

278 In Table 2, we present the number of neurons in our dataset, categorized by region. A
279 detailed overview of the differences in firing properties per auditory region is also provided in
280 **Figure S4-2** and **Supplementary Table 3**. Consistent with previous work, we found that firing
281 properties such as minimal latency differed between auditory regions in both age groups, with
282 AUDp and TEa being most distinct (**Fig. S4-2, Supplementary Table 3**) [22].

283 Notably, the effect size of the differences in baseline firing rates (FR) between groups
284 was larger in AUDd, AUDp, and AUDv. However, the evoked FR was smaller in all four areas.
285 The coefficient of variance of the FR was larger in AUDd, AUDp, and AUDv. Additionally,
286 the latency to peak was larger in AUDd and AUDp, while the full-width half-maximum of the
287 FR was also larger in these two areas. The minimal latency was larger in all four areas, and the
288 number of responsive trials was smaller in all four areas. Finally, the lifetime sparseness (see
289 Methods) was smaller in AUDp and AUDv between adolescent and adult neurons
290 (**Supplementary Table 3**).

291 Next, to study how adolescents and adults encode task performance we divided
292 neuronal responses by trial outcome (**Fig. 4d**). Specifically, we assessed neuronal
293 discriminability between stimulus-related and choice-related activity by calculating the area
294 under the curve (AUC) from a receiver operating curve (ROC) (**Fig. 4e**). Stimulus-related
295 activity was calculated as the difference between hit and false alarm trials in the easy and the
296 hard task separately. Choice-related activity was calculated as the average difference between
297 false-alarm and correct reject trials (miss trials were excluded since we had insufficient trial
298 numbers of this outcome). **Figure 4e** shows the average AUC values over time from an example
299 neuron (same neurons shown in **Fig. 4D**), and **figure 4f** for all neurons. The vast majority of
300 neurons successfully discriminated (AUC above shuffled data) either or both stimulus- and
301 choice-related activity (adolescent: easy stimuli = 93%, hard stimuli = 93%, choice= 95%;
302 adult: easy stimuli = 97%, hard stimuli = 97%, choice= 95%). The onset-latency of
303 discriminability was significantly slower, maximal discriminability significantly weaker, and
304 the duration of discriminability significantly shorter in adolescent neurons compared to adult
305 neurons (**Table 2**; plotted in **Fig. 4g**). We found no differences between auditory regions (**Fig.**

306 **S4-3**). Thus, auditory processing of stimulus- and choice-related activity in the adolescent ACx
307 has not yet reached full maturity.
308
309 *Population decoding of the hard task in adults outperforms adolescents*
310 While individual neurons exhibited weaker discriminability in adolescents than in adults, at a
311 population level these effects could saturate out. We therefore tested the ability of recorded
312 populations to decode trial outcomes. We used a linear discriminant analysis (LDA) decoder to
313 quantify hit versus correct reject trial outcomes. This comparison extracts information of the
314 combination of stimulus and choice (i.e. task-related differences). We decoded separately in
315 easy trials and hard trials, based on population activity during the first 200 milliseconds. For
316 the easy task there was no significant difference between decoding accuracy in adolescents and
317 adults (**Fig. 5a**). However, decoding accuracy was significantly higher in adults compared to
318 adolescents in hard trials (**Fig. 5a**). For both adults and adolescents, decoding accuracy was
319 significantly lower in hard trials compared to easy trials (**Fig. 5a**). We performed finer timescale
320 decoding to quantify the onset latency of differences between the activity for hit vs. correct
321 reject trials, defined as the time in which the decoding accuracy first crosses three times the
322 standard deviation away from the baseline (see Methods). We found that the latency was shorter
323 in adults in both the easy and hard trials (**Fig. 5b**). In adolescent animals the latency was larger
324 in hard trials (**Fig. 5b**) but not in adult animals (**Fig. 5b**). We then extended this analysis over
325 a temporal window from -0.5s to10s relative to tone onset (**Fig. 5c**). In the 200ms window after
326 the response period (2.0-2.2 s), there is no longer a significant difference in decoding accuracy
327 between adult and adolescent mice as well as between the two tasks (**Fig. S5-1**).
328 To further quantify the different decoding performance between two tasks, we plotted
329 the fisher separation metric ($\frac{\sigma^2_{between}}{\sigma^w_{within}}$, see Methods) for each level of difficulty. The separation
330 ratio of hard over easy tasks was significantly higher in adult mice than in adolescent mice
331 (robust linear regression, see Methods). Separation could have changed due to alterations in the
332 mean or the dispersion around the mean. Testing the dispersion, we found that the variance

333 around the means was not equal across easy and hard trials but rather significantly increased
334 for the hard trials (**Fig. 5d-f**). However, there was no significant difference in variance between
335 adolescent and adult mice within the same task **(Fig. 5d-f**).
336
337 *Effects of age and learning on cortical plasticity in mice engaged in the task*
338 We next studied how learning contributes to cortical plasticity in the different age groups. To
339 separate between age- and learning-related effects we recorded from the ACx of two new
340 groups of mice (adolescents and adults) that are novice on the task (**Fig. 6a**, P38, recordings=
341 6; P78, recordings =6, n = 3 mice per group). We collected 130 tone-responsive (by excitation)
342 neurons in novice adolescents and 186 tone-responsive (by excitation) neurons in novice adults
343 (**Fig. 6b-c**; **Supplementary Table 2**). Notably, novice mice did not discriminate between
344 sounds; instead, they licked indiscriminately to the tones presented during the task (**Fig. S61a-**
345 **b**).
346 To reveal learning-related differences, we compared novice and expert mice of similar
347 ages. We computed single-neuron AUCs for Go and No-Go trials (here we considered all trial
348 outcomes to reliably assess differences between novice and expert mice, since expert mice
349 mostly performed CR and novice mice mostly performed FA to No-Go tones), in both easy and
350 hard stimulus pairs (**Fig. 6d-g**). The onset-latency for AUC discrimination was significantly
351 different between adolescents and adults, but not between novice and expert mice, nor between
352 the easy and the hard task. The maximal AUC discriminability was higher, and the duration
353 longer in adults compared to adolescents, between experts and novice, and between the easy
354 and hard tasks (**Table 3**). These results suggest that plasticity is reflected in the interaction of
355 age and learning.
356 To compare neuronal discriminability and behavior more explicitly, we analyzed the
357 correlations between neuronal decoding and behavioral performance using linear regression. In
358 the easy task, both age groups showed high correlation between neuronal discriminability and
359 behavioral discriminability (**Fig. 6h-i**; light blue). In the hard task, only adult neurons had

360 significant correlations between neuronal discriminability and behavioral discriminability (**Fig.**
361 **6h-i**; dark blue).
362
363 *Effects of age and learning on tuning properties in passively listening mice*
364 Auditory discrimination learning has been shown to induce changes in the tuning properties of
365 sounds in the ACx [18, 19, 21, 23-28]. To study whether learning-induced plasticity in ACx is
366 distinct in adolescents and adults, we recorded from the same mice described above under
367 passive listening condition, by simply extending the recording following the engaged session
368 (**Fig. 7a, Supplementary Table 5**; Adolescents recordings = 4; Adults recordings = 4). These
369 data were compared to a group of novice mice, recorded under passive listening following their
370 engaged session (**Supplementary Table 5**; adolescent recordings =6; adult recordings =6).
371 We characterized the response profile of ACx neurons using a 'frequency-response
372 area' protocol composed of twenty pure tones (4-40 kHz, spaced at 0.1661 octaves steps), each
373 played for 100ms at five different attenuations (72-42 dB SPL). We limited our analysis to
374 significant excitatory units only, determined as being auditory responsive in a window from
375 tone onset to 50ms after tone offset at 62 dB SPL. Our dataset included four groups. The two
376 adolescent groups were: 1) adolescent novice, and 2) adolescent experts, with 92 and 80
377 neurons, respectively. The two adult groups were: 3) adult novices, and 4) adult experts, with
378 123 and 84 neurons, respectively. Neurons in all groups responded to pure tone as expected
379 from previous studies in adult mice, including classical V-shaped frequency response areas [22,
380 29-32]. A representative example from a responsive neuron of the adolescent expert group and
381 one from the adult expert group are shown in **Figure 7b-c**. We next compared the firing
382 properties of different sub-regions (TEa and AUDd were excluded since less than 20 units
383 responded to the FRA protocol after task engagement, see **Supplementary Table 5**). We found
384 that the firing properties were distinct between adolescent and adult neurons, in both AUDp
385 and AUDv. However there were no significant differences between AUDp and AUDv in either
386 group (**Supplementary Table 6**). The peak of frequency tuning was heterogeneous across the
387 frequency range with no clear trends of learning or age (**Fig. 7d**). To quantify these responses,

388 we calculated tuning width (**Fig. 7e**), population sparseness (**Fig. 7f**), the distance between the
389 best frequency of the neurons and the Go frequency of the easy task (**Fig. 7g**), and single neuron
390 discriminability (**Fig. 7h**). As expected from previous work, learning significantly changed
391 neuronal tuning properties in adults (**Fig. 7e-h**, 'Adults'). Surprisingly, however, learning had
392 no significant effects on any of the single neuron parameters that we tested in the adolescent
393 group (**Fig. 7e-h**, 'Adolescents' In adults, these results remained consistent when AUDp and
394 AUDv were analyzed separately (**Fig. S7-1**). In contrast, adolescent results were more
395 heterogeneous: AUDp exhibited changes in tuning alongside higher discriminability, while
396 AUDv showed no detectable differences. Thus, despite the notion that adolescents' brains may
397 be more plastic than adults, we found no clear evidence for learning-induced plasticity in basic
398 tuning properties of the adolescent mice following our protocol (**Table 5, Fig. S6-1**). These
399 results underscore that learning-induced plasticity manifests differently in adolescents and
400 adults.
401

402 # Discussion

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

411

412 *Adolescence behavior – perceptual and cognitive differences*

413 We utilized the Educage platform to compare between adolescents and adults[18]. The Educage
414 provides an optimal environment for training young mice; its automated nature minimizes
415 human handling, which increases efficiency and reduces to a minimum any interference during
416 the training process. Using the Educage, we found that procedural learning was intact in pre-
417 adolescent mice (P20-P25), and behavioral differences were found only after the mice learned
418 the task during the discrimination phase (P30-37). Compared to adults, adolescent mice scored
419 lower on the hard versions of the task (**Fig. 1i-j**). Interestingly, our data suggests that the lower
420 performance in the hard task in adolescents was not due to a deficiency in perceptual sensitivity
421 (**Fig. 2b**), but rather to a cognitive deficit – namely, response bias (**Fig. 2c**). This result is
422 perhaps not surprising since higher response bias is a cognitive phenomenon that has been well
423 documented in adolescence as a critical period of inhibitory control[36]. Indeed, additional
424 cognitive control mechanisms are still developing during this age [37-39].

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

429 which we did not manipulate[41]. Notably, few studies have explicitly characterized auditory
430 discrimination of adolescents in rodents; the majority of these have been carried out in gerbils
431 using an amplitude modulation (AM) detection task [4, 26, 42]. Despite the different task, many
432 of our age-related findings are aligned with the findings in gerbils, including lower performance
433 in adolescents, and greater variance in adolescent behavior[4]. However, some findings differ.
434 For example, the response bias which was different in our data, was similar between young and
435 adult gerbils in an AM modulation detection task. This difference may arise from differences
436 in task design, as well as the severity of the punishment used in the task (here, we used a milder
437 punishment). Together, our data support a view that cognitive factors, and to a lesser extent
438 perception of pure tones *per se*, are the primary contributors to the poorer performance observed
439 in adolescent mice when performing a sound discrimination task.
440
441 *Representation of stimulus and choice in ACx still undergoes maturation in adolescence*
442 Motivated by the abovementioned behavioral differences, we tested how single neurons in
443 ACx—whose involvement in task performance we confirmed via optogenetic silencing in adult
444 mice (**Fig. S3-2**)—represented sounds (by comparing Hits vs FA) and choices (by comparing
445 FA to CR). While some studies suggest that ACx is not required for simple auditory-guided
446 behaviors [43], others highlight its role in pure tone perception and task execution [19, 20, 44,
447 45]. While cortical representations of sounds in ACx could be the basis of perceptual
448 constraints, distinct processing of choices would reflect that more cognitive mechanisms are
449 involved. We found that cortical representations of sounds and choices in mice engaged in the
450 task were still immature as compared to adults, e.g., responses in adolescents were more
451 sluggish and less informative (**Figs. 4-6**). These differences can be a result of age-related
452 differences, and/or the result of a combined 'age x learning' effect.
453 We found modest age-related differences in spiking activity of neurons from layer 5/6
454 between adolescents and adults (**Fig. 6d-g; Fig. 7**). This limited difference is perhaps not
455 surprising because the CP for tonal organization in ACx is already complete by the age at which
456 we tested the animals (i.e., on P38 [9, 13]). In contrast, more pronounced differences emerged

457 from the interaction of 'age x learning', specifically in mice that learned the task and were
458 actively engaged in the behavior (**Fig. 4e, f; Fig. 6d-g**). This interaction is evident in the reduced
459 single-neuron discriminability (AUC) during the hard task for expert adolescent mice (**Fig. 6d**
460 **vs. 6f**). While this may seem counterintuitive, it is consistent with the idea that learning can
461 drive complex reorganization in sensory representations. For example, we recently showed that
462 learning leads to distinct and sometimes opposing effects on single-neuron responses with
463 respect to whether the task is perceptually easy or difficult [46].

464 It is important to note that our recordings were restricted to infragranular layers (5/6)
465 and did not sample supragranular layers (e.g., layers 2/3), which may exhibit distinct plasticity
466 profiles [21]. As such, the developmental differences we report may reflect layer-specific
467 properties of deep output layers, particularly their involvement in learning-related plasticity and
468 communication with subcortical or frontal targets. We speculate that immature feedback (FB)
469 to ACx may underlie these differences for several reasons. First, FB from higher cortical regions
470 to more primary cortices has been proposed as a neural substrate for difficult perceptual
471 discriminations[47, 48]. Second, FB activity (often referred to as 'top down') is thought to be
472 the key neural pathway involved in cognitive control mechanisms[49]. FB is typically
473 expressed in the late (>100ms) phase of single neurons activity, which was significantly
474 different between adults and adolescents in our data (**Fig. 4d**). Furthermore, FB may act to
475 stabilize or suppress sensory responses in ACx, reducing apparent single-neuron
476 discriminability of hard stimuli after learning (**Fig. 6f**), while preserving or enhancing in the
477 population discriminability such as we observed in **Fig. 5**.

478 The precise source of FB that contributes to perceptual or cognitive differences remains
479 to be determined, but likely candidates include regions that project to ACx [50, 51], such as
480 medial prefrontal cortex[52, 53], orbitofrontal cortex[54], or anterior cingulate[55]. Testing the
481 causal role of these regions during adolescence will be an important direction for future
482 research.

483 Cortical coding is performed by populations of neurons, which can efficiently encode
484 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**). This result suggests that the cortical network uses shared representation for decoding the easy and hard tasks, possibly by the same neurons. This is consistent with our recent work showing that indeed the same neurons in ACx respond to the same stimulus differently, depending on if the mouse is engaged in discriminating an easy sound pair or a hard sound pair[46]. The shallower slopes in adolescence suggest a less efficient decoding scheme that may arise from different populations of neurons encoding the stimuli in the easy and hard versions of the task.

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

513 *Learning in the adolescent brain*

514 Cortical representations of pure tones have a critical period (CP) between P12-P15[13, 16, 17].
515 Indeed, we chose to study animals at P38-P40 on a pure tone discrimination task precisely to
516 avoid the hypersensitivity associated with this CP. But this may be unavoidable, as recent work
517 has shown that ACx also has a later CP for more complex sounds features[16, 17]. Those late
518 CP's overlap with our training schedule and the underlying biological manifestation of this CP
519 is expected to affect learning related plasticity in general. The underlying mechanism of the late
520 CP to complex sounds has been associated with dynamic changes of inhibitory circuits[62, 63].
521 Thus, immature inhibitory circuits remain sensitive in adolescents, and could affect how pure
522 tones are represented, despite the closure of the classic tonal CP. Notably, sound discrimination
523 learning changes tonal representations even in adult mice, well after the closure of CP's for
524 both simple and complex sounds [18, 24, 46]. We hypothesize, therefore, that the differences
525 in performance on more difficult discriminations between adolescents and adults are shaped
526 primarily by age-specific learning-related plasticity, rather than by differences in pure tone
527 representation per se.

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

539

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, S3b, 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.

561 **Supplemental information:**

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

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

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

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

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

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

568 S4-2 *Adult and adolescent firing properties are distinct between auditory regions during task*

569 *engagement,* related to Figure 4.

570 S4-3 *The neuronal discriminability of stimulus- and choice related activity is similar across*

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

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

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

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

575 *auditory cortex*, related to Figure 7.

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

577 related to Figure 3.

578 Supplementary Table 2 *Adolescent and adult firing properties are distinct between different*

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

580 Supplementary Table 3 *Adolescent and adult mice have distinct firing properties in different*

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

582 Supplementary Table 4 *Neuronal statistics during passive FRA protocol in expert and novice*

583 *mice,* related to Figure 4.

584 Supplementary Table 6 *Adolescent and adult mice have distinct firing properties across*

585 *different sub-regions of ACx—passive listening*, related to Figure 7.

586

587

588

589 # Methods

590 ***Animals***

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

597

598 ***Surgical procedures***

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

601

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

607

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

616 catalog #KWIK CAST). After the craniotomy, mice were injected with 0.04mg/g of 10%
617 Meloxicam solution and 0.2 ml Saline. During the 48 hours of recovery, mice received two
618 further doses of 0.04mg/g of 10% Meloxicam solution. 24h before recordings, mice were again
619 anesthetized to remove regrowth of the dura-mater during training. The area around the
620 craniotomy was cleaned with Hydrogen-peroxide (3%), the silicon elastomer was replaced, and
621 another dose of 10% Meloxicam was injected.
622
623 *Virus Injection & Cannula Implantation.* The surgical procedure was similar to the *head-bar*
624 *implantation & craniotomy*. The differences in the procedure are outlined below. Bilateral
625 craniotomies were performed at 2.5 mm posterior and 4.3 mm lateral to the bregma. We injected
626 200nl AAV5-CAMKII-GtACR2-FRED_kv_2.1, or 200nl AAV9-CAMKII-dTomato. Both
627 viruses were produced at the ELSC virus-core facility. We injected the virus at a depth of 1.1
628 mm at 0°. Afterwards, we inserted a 0.2 mm diameter optical fiber with an attached cannula
629 (CFML1202; Thorlabs) at a depth of 0.9 mm at 0°, bilaterally. We chronically fixed the fiber
630 position on the skull using dental cement (Meta bond).
631
632 ***Auditory Stimuli***
633 Sound stimuli were delivered from a calibrated free-field speaker (ES1 SN:4568) using a
634 multifunction processor (ED1; Tucker-Davis Technologies). The speakers were calibrated with
635 a free-field microphone (Type 4939, Bruel and Kjaer). The stimuli were comprised of pure
636 tones, sampled at 500 kHz, between 7.07 to 14 kHz for behavioral training and 4-40 kHz for
637 passive listening. All stimuli were played at sound pressure levels of 72 dB SPL for behavioral
638 training, and 72-, 62-, 52-, and 42-dB SPLs for passive listening. All tones were played in
639 randomized order.
640
641 ***Behavior***
642 *Automated home-cage training (the 'Educage').* For a detailed description of the EduCage
643 behavioral platform, see [18]. After RFID- implantation, groups of 5 mice were placed in their

644 home-cage that was connected to a behavioral chamber called the Educage (Fig. 1a). Mice were
645 provided with food *ad libitum*. Water access was restricted to the Educage chamber. Mice could
646 access the Educage freely to retrieve water mixed with 5% sucrose. The behavioral training was
647 controlled by a custom LabView 2019 (National Instruments) program running on BNC-2110
648 (National Instruments) and FPGA (MyRIO – National Instruments) to register licks, RFIDs
649 numbers, and deliver stimuli and reinforcements. Every time a mouse entered the water port its
650 RFID was automatically recognized, a trial initiated, and an auditory stimulus played from a
651 speaker vertically positioned above the behavioral chamber. We trained mice gradually for
652 three weeks on a Go/No-Go paradigm of pure tones, spaced around a category boundary of 10
653 kHz. This paradigm enabled us to efficiently compare between adolescent and adult behavior,
654 since learning was restricted from post-weaning (P20) to adolescence (P37) (correspondingly
655 adults were trained from P60-P77). Mice were trained in four separate training phases (Fig.
656 1B). For the first 24h (P20-P21 in the adolescent group of mice and P60-P61 in adult mice)
657 mice were habituated to the behavioral chamber and could approach the port freely to retrieve
658 water after licking the waterspout (0.0015 ml per trial). During the following 48h (P21-P23 in
659 the adolescent group of mice and P61-P63 in adult mice) mice underwent a tone-association
660 phase. All tones played were 300ms long (sampled at 500 kHz). Tone association was restricted
661 to the lowest Go tone (7.07 kHz at 100% probability). Mice received a water-reward in response
662 to licking during the first 2 seconds after tone-offset. This was followed by the first tone-
663 discrimination stage, which contained a No-Go tone (14.14 kHz at 45% probability) at 1 octave
664 away from the Go tone. We analyzed four response outcomes—Hits, Misses, Correct
665 Rejections, False Alarms— in the 2 seconds response window after the tone-offset. 'Hits' were
666 counted after 5 successful licks to the Go-tone and followed by a water-reward. 'Misses' were
667 counted when licks to the Go-tone trials did not pass the lick threshold. Correspondingly,
668 'Correct Reject' (CR) trials, were all No-Go trials below the lick threshold. CR trials were not
669 reinforced. 'False Alarms' (FA) were counted when 5 or more licks were registered after the
670 No-Go. FA trials were followed by a white-noise punishment (5-20 kHz at 72 dB SPLs, 2
671 seconds) and an inter trial delay of 5 seconds. In addition, mice were exposed to 5 probe trials

672 (tones whose trial outcome that were neither rewarded nor punished at 8.49 kHz, 9.567 kHz,
673 10 kHz, 10.44 kHz, and 11.89 kHz; played at 2% per tone). Mice trained on the 1 octave tone
674 discrimination for 1 week (P23-P30 in adolescent mice and P63-P70 in adult mice). Next, we
675 added a second pair of tone - a Go/No-Go tone pair spaced 0.25 octaves apart (9.17 kHz and
676 10.95 kHz) for another week of training. All tones were played in randomized order and trial
677 difficulties of the easy and the hard task were intermingled. The experiment was terminated at
678 age P37 for adolescent and P77 for adult mice.
679
680 *Head-fixed training*. We used a head-fixed paradigm that enabled us to efficiently train
681 adolescent and adult mice on the same task and record from engaged mice after learning. We
682 trained mice on the same Go/No-Go stimulus setup of the Educage described above with few
683 differences which are outlined below.
684 The head-fixed behavioral training was controlled by a custom MATLAB 2023B
685 (Mathworks) program running on BNC-2110 (National Instruments) to register licks and
686 deliver stimuli and reinforcements. Mice were water-restricted for 24h prior to training, and
687 their weight monitored daily. While training on the setup, mice received water supplemented
688 with 5% sucrose, depending on their weight (0.125 ml per g of body weight). Additional water,
689 in the form of Hydrogel (ClearH2O) was given after daily training if mice did not consume the
690 required amount during training, or their body weight dropped below 85% of the pre-restriction
691 weight. Mice were gradually habituated to the setup by head-fixation and received free water
692 after licking, during the first training day (P23 in adolescent mice and P63in adult mice).
693 Licking was measured with an optical lick-meter (Sanworks). During the second training day
694 mice were associated with a tone. All tones of the task were 100ms long (sampled at 500 kHz).
695 Mice received a water reward in response to licking in the first 2 seconds after tone offset.
696 Reward-reinforcement was delayed to 0.5 seconds after the tone offset. To break the regularity
697 of the trial sequences, we applied a dynamic inter-trial interval of 6-8 seconds. New trials were
698 initiated if mice did not lick the spout at least 2 seconds prior to the upcoming trial. Mice were
699 familiarized with the tone for at least two training days (P24-26 in adolescent mice and P64-66

700 in adult mice). Then, and when the lick rate was above 80%, mice proceeded to the first tone
701 discrimination stage. We gradually increased the probability of No-Go trials, depending on the
702 lick rate of the mouse. We added the 0.25 octave Go/No-Go tone pair, after a maximal of seven
703 training days on the 1 octave tone pair (P26-33 in adolescent mice and P66-73 in adult mice),
704 or if the behavioral performance exceeded a threshold of d' = 1, as calculated using the signal
705 detection metric. Similar to the Educage, all tones were played in randomized order and trial
706 difficulties of the easy and the hard task were intermingled. Mice were trained until P37 in the
707 adolescent group and P77 in the adult group.
708
709 ***Optogenetic Manipulation***
710 Adult mice (P60) were injected with AAV5-CAMKII-GtACR2-FRED_kv_2.1 (n = 3) and
711 implanted with cannulas prior to training as described above. Mice underwent optogenetic
712 stimulation after training on the head-fixed task by attaching a ferrule patch cable (M79L01;
713 Thorlabs) to the implanted cannula using a mating sleeve (ADAF1; Thorlabs). The ferrule cable
714 end (SMA connector) was connected to an LED driver setup (LEDD1B; Thorlabs) that enabled
715 us to send precise outputs at 476 nm (blue light; 5mW) bilaterally. To test the effect of ACx
716 inhibition on task performance, we timed the LED output to 50ms before tone onset until 50ms
717 after tone-offset. The LED consisted of pulses given at 10 Hz for 200ms
718 Behavioral sessions of optogenetic manipulation were identical to the Go/No-Go
719 paradigm of previous training sessions. We inhibited ACx in 50% of the trials in randomized
720 order. Optogenetic manipulations were repeated throughout multiple sessions of expert
721 performance. As a control, we repeated the same experiment in mice injected with AAV9-
722 CAMKII-dTomato (n = 3). After completion of the experiment, animals were deeply
723 anesthetized with Ketamine-and medetomidine (0.008 g/kg, and 0.65 mg/kg, respectively) and
724 then perfused with 4% Paraformaldehyde. After the perfusion, the brain was extracted and
725 preserved. We then sectioned the ACx using a cryostat (Leica) into 0.05 mm thick coronal
726 slices. Afterwards brain slices were washed one time under 1% Phosphate-buffered saline
727 (PBS), and a second time with 1% PBS plus 0.4% Triton. Finally, slides were mounted and

728 stained with DAPI (4%) and imaged using a wide-field fluorescent microscope (Olympus Life
729 Science, Olympus IX83), to reconstruct the probe position offline.
730
731 ***Extracellular Recordings***
732 *Recording setup.* Before the onset of the recording session, we removed the silicon elastomer
733 and placed an external reference electrode (Ag/AgCl wire) on the skull of the right hemisphere.
734 All recordings were performed using Neuropixels 1.0 (Npx; IMEC, phase 3A), together with a
735 base-station connected to a chassis (IMEC; NI PXIe -1071, National Instruments). Probes were
736 mounted to a custom-made steel rod and connected to the ground. Before penetration, probes
737 were covered with a fluorescent dye (Dil (Invitrogen catalog #V22885- red, or Dio (Invitrogen
738 catalog #V22886 – yellow), to reconstruct penetration sites after the recording.
739 Then, Npx probes were inserted into the left ACx at 30° to a depth of 3.85 mm (this approach
740 limited our recording to deep layers of the ACx). The skull surface was submerged in saline
741 and the probe allowed to settle for 10 min. We sampled recordings at 20 kHz, with action
742 potential band filtered to contain 0.3- to 10kHz frequencies. Action potential band gain was set
743 to 500. Out of the 960 available sites on the 1 cm shank of the Npx probe [68], we acquired the
744 384 lowest recording shanks in a staggered configuration. We used common-average
745 referencing to process the voltage-traces.
746
747 *Recording Schedule.* We performed recordings of both novice and expert adolescent and adult
748 mice. Novice recordings were performed after tone-association (see *head-fixed training;*
749 adolescent n=3; adult n = 3). Expert recordings were performed after learning (adolescent n=
750 5; adult n = 6). Both novice and expert recording include the easy (1 octave) and hard (0.25
751 octave) tone pair. Both recordings were performed after P37 for adolescent mice and P77 for
752 adult mice. Tone probabilities and protocol parameters were identical to the *head-fixed training*.
753 Each mouse was recorded multiple times (novice recordings adolescents n = 6; novice recording
754 adults n =6; expert recordings in adolescents n = 13; expert recordings in adults n = 14). In
755 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 = 5; novice recording adults n = 5; 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). In all figures, unless mentioned otherwise, we represent mean and median of the

784 data, together with standard deviation, quartile, or standard error of the mean (MAD/sqrt(n); n
785 = number of data points).
786
787 *Behavioral analysis.* We categorized behavioral responses of the task into hit, miss, false alarm,
788 and correct reject responses. To compare behavioral performance, we calculated the d' value
789 based on the signal detection metric (d' = the inverse normal distribution of the z-scored hit
790 rate) – the z-scored (false alarm rate)). Z-scored values of 0 were rounded to 0.01 and z-scored
791 values of 1 were rounded to 0.99, to avoid d' approaching infinity. Next, we calculated
792 psychometric curves based on the lick rate (above the lick threshold) to learned and probe trials
793 (Fig. 2b, c, e). Psychometric curves were normalized and fitted to a sigmoidal function, defined
794 as follows:
795
796 $$S(t) = a\frac{1}{1+e^{\frac{-(t-b)}{c}}}$$ (Equation 1)
797
798
799 Where *a* denotes the lick rate, *b* = the time of the inflection point, and *c* = the steepness of the
800 curve. Based on the normalized psychometric curve, we extracted two inflection points (random
801 lick rate at 0.5 and category boundary at 10 kHz). To compare the heterogeneity of adolescent
802 and adult behavior, we calculated the criterion bias (the inverse normal distribution of the (-
803 0.5*(z-scored (hit rate) + z-scored (false alarm rate))) and analyzed the coefficient of variation
804 (CV) of behavior across the learning. We applied similar behavioral measurements to the
805 Educage training and head-fixed training, as well as recordings and optogenetic manipulation
806 during task performance. During expert-performance in the head-fixed task, we analyzed the
807 running d' as the average performance of the last 25 trials (**Fig. 3**). In addition, we also
808 calculated the average lick trace during the reinforcement delay. Here, we also extracted the
809 lick latency per trial per mouse, defined as the lick latency from tone-onset (**Fig. 3**).
810

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:

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

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

837 *Single neuron analysis during task performance.* We investigated the FR of excitatory auditory
838 responses per trial from 200ms preceding the stimulus upto 600ms after the tone onset. To test
839 how well single neurons discriminate between trial outcomes, as well as task difficulty, we
840 calculated task-related activity per trial using receiver operating characteristics (ROC) and
841 calculated the area under the curve (AUC) across a running window of 25ms in bins of 50ms
842 (Fig. 4, Fig. 6). Stimulus-related activity was defined as the hit compared to the false alarm
843 trials, separately for the 1 octave and 0.25 octave tone pair. Stimulus-related activity of miss
844 compared to correct trials were not analyzed due to the lack of miss trials in expert performance.
845 In addition, we also tested the average (mean of 1 octave and 0.25 octave tone pair) choice
846 discrimination, defined as the difference between false alarm and correct reject trials (Fig. 4).
847 We sampled 20 trials per trial outcome and repeated the ROC encoding 100 times. Afterwards,
848 we adjusted the average AUC values of all iterations of each neuron to its baseline AUC before
849 the tone onset ((AUC – baseline AUC) + 0.5). AUC values close to 0.5 indicate a low neuronal
850 discriminability, and values upto 1 indicate high neuronal discriminability [72]. To test
851 significant discrimination, we repeated the ROC encoding with shuffled trial identities (20
852 randomly shuffled trials per trial outcomes across 100 iterations). Significant discriminability
853 of a neuron was defined as the AUC time bins that exceed the mean ±3 std of the shuffled
854 distribution [64]. Adolescent and adult neuronal discriminability were compared using the
855 onset-latency of discrimination of a neuron (first timepoint > mean ±3 std of the shuffled
856 distribution; Fig. 4h, Fig. 6c), the maximal AUC of the running window (Fig. 4j, Fig. 6e) and
857 the duration of discriminability (time duration > mean ±3 std of the shuffled distribution; Fig.
858 4i, Fig. 6d). To compare novice and expert performance we repeated the AUC encoding and
859 compared the general discrimination of 1 octave and 0.25 octave Go and No-Go discrimination
860 of all trial outcomes (Fig. 6). To evaluate the correlation of behavioral performance and
861 neuronal discriminability, we compared d' per recording to the average of the maximal AUC
862 of all neurons per recording. The AUC and d' values were compared via pairwise Pearson
863 correlation (Fig. 6f).
864

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 and tasks. Data were standardized per neuron before computing variance per stimulus (Fig. 5e). To visualize the results, we projected the population representations to the space spanned by

892 the two LDA decoding dimensions for the easy and hard tasks. We averaged the mean and
893 covariance matrix across sessions for each stimulus and age group, representing the
894 distributions with corresponding ellipses (Fig. 5f).
895
896 *Single neuron analysis during passive listening.* We extracted the frequency-response areas
897 (FRAs; Fig. 7b) based on the pure-tone responses (from tone onset up to 50ms after tone-offset)
898 of auditory responsive neurons. The bandwidth of excited units was computed for significant
899 excitatory responses of adjacent frequencies at 62 dB SPL and subtracted by the number of
900 expected false-positive responses. Afterwards the sum of significant frequency responses per
901 neuron was multiplied by the octave distance between every frequency to receive an octave-
902 based bandwidth measure. We then calculated the population sparseness as the fraction of
903 significant excited responses in adolescent and adult novice and expert mice (Fig. 7f). To test
904 the neuronal discriminability to pure tone responses, we calculated the pairwise d' based on the
905 FR per trial in all frequencies at 62 dB SPL [22, 66]. For two given frequencies p and q, d'
906 values were calculated in the following way:

907 $$\frac{d(\mu_{p,}\mu_q)}{\mu(\sqrt{\sum_n^{i=1}(\widehat{p_i}-\widehat{q_i})^2})} \qquad \text{(Equation 4)}$$

908
909 Where d represents the distance between the mean FR (μ) of frequency p and q, as vectors with
910 n entries representing the mean signal (averaged over individual trials) in n-dimensional space
911 for frequency p/q. The average distance (d) is then divided by the mean of the inner Euclidian
912 distance ($\sqrt{\sum_n^{i=1}(\widehat{p_i}-\widehat{q_i})}$) between each single trial (Fig. 7h).
913
914 *Statistical Analysis.* Statistical comparisons were performed in custom-written codes in
915 MATLAB 2023b (MathWorks) and Python 3.12. We assessed if the data was normally
916 distributed using a Kolmogorov-Smirnov test. Normally distributed data was tested using a
917 paired (within group comparison) or independent (between group comparison) t-test and

presented with mean ± SEM. Non-normally distributed data was tested using a Wilcoxon sign rank (within group comparison), or rank-sum (between group comparison) test. Multiple-samples were tested with ANOVA (parametric data) and Kruskal-Wallis tests (non-parametric data). All tests corrected after multiple comparisons (Bonferroni-correction for two-samples and Tukey-Kramer for multiple samples). In addition, we applied linear-mixed-effect models (LME) to account for hierarchical data structures and variability within groups of co-housed mice (Table 1), repeated measurements of mice and recordings (supplementary Table 1, Table 3, Table 4), as random effects. Fixed effects, and pairwise interaction effects were adjusted with post-hoc tests and multiple comparisons. For all statistical tests significant differences were defined as p-values below 0.05.

929

930 Figure Legends

931 **Figure 1:** *Adolescent mice exhibit lower performance during self-initiated auditory learning in*
932 *the 'Educage'.* **A**. Schematic model of the 'Educage' (left), the trial structure and trial types
933 (FA, false alarm; CR, correct reject). **B.** Experimental timeline. Total training time was 21 days
934 (±2). **C.** The sounds used for training. Light blue- easy task; dark blue- hard task; Grey- probe-
935 trials. **D.** Learning curve examples. Adolescent mouse (left, P20-to-P37); adult mouse (right,
936 P60-to-P77). Vertical dashed lines indicate the easy-hard transition. Horizontal line is d'=1. **E.**
937 Number of trials to reach threshold (d'>=1; adolescents, n =15; adults, n=15; z=-0.1659;
938 p=0.8682, two-sample Wilcoxon rank sum test). **F.** Discriminability (d') of the easy task in
939 adolescent mice at P30 (grey; n=15, d'=2.0001±0.1791; n trials=8317±712) and adult mice at
940 P70 (black; n=15, d'=2.1986 ± 0.2441; n trials=13229 ± 514, top: marked by the arrowhead).
941 Dashed lines: mean trials per group (t-stat=-5.6314, df=28 ,p=4.9566e-06, two-sample
942 independent t-test, vertical line), and mean d' per age group (z=-0.2074; p=0.8357, two-sample
943 Wilcoxon rank sum test, horizontal line) **G.** Change in discriminability (Δd') of the easy task
944 before and after the introduction of the hard task (Top: arrowheads; left: adolescent, signed
945 rank=14, p=0.0067; right: adult, signed rank=14, p=0.1514, one-sample Wilcoxon signed rank
946 test; Δd' between adult and adolescent mice: z=-2.0739; p=0.0381, two-sample Wilcoxon rank
947 sum test) **H.** Same as 'G' for the first 100 and last 100 trials of the experiment in the easy task
948 (adolescent signed rank=120, p=6.1035e-05; adult signed rank=120, p=6.1035e-05, one-sample
949 Wilcoxon signed rank test; Δd' between adult and adolescent mice: z=-1.0370; p=0.2998, two-
950 sample Wilcoxon rank sum test). **I.** Same as 'F' for the hard task (adolescent-grey; n=15,
951 d'=1.1895±0.1783; n trials=4163 ± 297; adult-black; n=15, d'=1.8342 ± 0.1743; n
952 trials=4102±475) (mean trials per group: t-stat=0.1306, df=28, p=0.8970, two-sample
953 independent t-test, vertical line; mean d' per group: z=-2.2398; p=0.0251, two-sample
954 Wilcoxon rank sum test, horizontal line). **J.** Same as 'G' for the hard task. (adolescent signed
955 rank=73, p=0.4887; adult signed rank=114, p=8.5449e-04, one-sample Wilcoxon signed rank
956 test; Δd' between adult and adolescent mice: z=-1.9495; p=0.0512, two-sample Wilcoxon rank
957 sum test).
958
959

960 **Figure 2***: Adolescent mice exhibit lower response inhibition and higher variability in*
961 *performance over the course of learning.* **A.** Schematic timeline of tone discrimination). **B.**
962 Average psychometric curve of the proportion of lick responses (above the threshold) across
963 all stimuli (probe and learned; No-Go stimuli: p = 0. 0085, Go stimuli: p = 0. 0963, Wilcoxon
964 rank sum test, after Bonferroni correction). **C.** Same as 'B',but normalized and fitted to a
965 sigmoid function0(proportion to the lick threshold at 10kHz, p = 0.1163; frequency at 50%
966 proportion to the lick threshold, p = 0.5896, Wilcoxon rank sum-test, after Bonferroni
967 correction). **D.** Maximal lick bias (i.e., criterion bias) per mouse during the last 2000 trials (c-
968 bias: p = 0.0136, Wilcoxon rank sum test). **E.** Average psychometric curve, fitted to a sigmoid
969 function across all stimuli (probe and learned) at different periods of the experiment — Left:
970 first tertile-easy task (per mouse; adolescent n = 15, adult n = 15); Middle-left: last tertile – easy
971 task; Middle-right: first tertile-hard task; Right: last tertile-hard task. (Dashed vertical
972 =category boundary). **F.** Same as E. as the maximal C-bias. From left to right:(z = -0.5104, p =
973 0.6098, two-sample Wilcoxon rank sum test), (z = -2.6546, p = 0.0079, two-sample Wilcoxon
974 rank sum test), (z = -4.0739, p = 4.6228e-05, two-sample Wilcoxon rank sum test), (z = -3.2768,
975 p = 0.0011, two-sample Wilcoxon rank sum test). **G.** Change in lick bias (Δ c-bias) between
976 the last 100 trials before and after the introduction of the hard task. (left, adolescent signed rank
977 = 17, p = 0.0125; right, adult signed rank = 83, p = 0.2078, one-sample Wilcoxon signed rank
978 test; Δ c-bias between adolescents and adults, z = -2.9035; p = 0.0037, two-sample Wilcoxon
979 rank sum test). **H.** Coefficient of variation (CV) of the d' of the easy task before the introduction
980 of the hard task. (z = -0.1659, p = 0.8682, two-sample Wilcoxon rank sum test) **I.** Same as 'H'
981 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).

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; 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 and lick rate (LR) 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 pertrial 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 (+- SEM) 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 = 107; grey) and adult (n = 177; 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: *Unlike adults, neuronal tuning properties of adolescents do not change 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 (first 1000 trials) compared to "expert" (last 1000 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 per (z = 1.3349, p = 0.1819, Wilcoxon rank sum test). **D.** Average number of licks during ITIs that mice licked (z = -0.8007, p = 0.4233, 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 (CAMKII; bottom). **D.** Lick rate 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 (CAMKII, 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 Rate 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 (CAMKII, 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 (CAMKII 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 10 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: *Adult and adolescent firing properties are distinct between auditory regions during task engagement* **A.** Adolescent average baseline FR of excited single neurons in AUDd (blue), AUDp (purple), AUDv (magenta) and TEa (green) from tone-onset to 50ms after tone offset across all stimuli. Average baseline FR (Hz, statistical comparisons are presented in supplementary Table 2) **B.** Same as A. for the evoked FR (Hz) **C.** Same as A. for coefficient of variance of FR **D.** Same as A. for the latency to peak of maximal FR (ms).**E.** Same as A. for the full-width-half maximum of peak FR (ms) **F.** Same as A. for the minimal latency of peak FR (ms). **G.** Same as A. for proportion of responsive trials. **H.** Same as A. for the lifetime sparseness. **I-P.** Same as A-H. for adult neurons.

Supplementary Figure 4-3*: The neuronal discriminability of stimulus- and choice related activity is similar across auditory sub-regions.* 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 (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), 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 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 (fraction of tone-excited 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 firing properties 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 4: *Adolescent and adult 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 5: *Neuronal statistics during passive FRA protocol in expert and novice mice.* Acquired single units, acquired tone-modulated units, and fraction 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 6: *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 7: *Overview of analyzed datasets.* 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.

1382 17. Nakamura M, Valerio P, Bhumika S, Barkat TR. Sequential Organization of Critical
1383 Periods in the Mouse Auditory System. Cell Rep. 2020;32(8):108070. doi:
1384 10.1016/j.celrep.2020.108070. PubMed PMID: 32846128.
1385 18. Maor I, Shwartz-Ziv R, Feigin L, Elyada Y, Sompolinsky H, Mizrahi A. Neural Correlates
1386 of Learning Pure Tones or Natural Sounds in the Auditory Cortex. Front Neural Circuits.
1387 2019;13:82. Epub 20200128. doi: 10.3389/fncir.2019.00082. PubMed PMID: 32047424;
1388 PubMed Central PMCID: PMCPMC6997498.
1389 19. O'Sullivan C, Weible AP, Wehr M. Auditory Cortex Contributes to Discrimination of
1390 Pure Tones. eNeuro. 2019;6(5). Epub 20191016. doi: 10.1523/ENEURO.0340-19.2019.
1391 PubMed PMID: 31591138; PubMed Central PMCID: PMCPMC6795560.
1392 20. Ceballo S, Piwkowska Z, Bourg J, Daret A, Bathellier B. Targeted cortical manipulation
1393 of auditory perception. Neuron. 2019;104(6):1168-79. e5.
1394 21. Kanold PO, Nelken I, Polley DB. Local versus global scales of organization in auditory
1395 cortex. Trends Neurosci. 2014;37(9):502-10. Epub 20140704. doi:
1396 10.1016/j.tins.2014.06.003. PubMed PMID: 25002236; PubMed Central PMCID:
1397 PMCPMC4152386.
1398 22. Feigin L, Tasaka G, Maor I, Mizrahi A. Sparse Coding in Temporal Association Cortex
1399 Improves Complex Sound Discriminability. J Neurosci. 2021;41(33):7048-64. Epub 20210708.
1400 doi: 10.1523/JNEUROSCI.3167-20.2021. PubMed PMID: 34244361; PubMed Central PMCID:
1401 PMCPMC8372025.
1402 23. Bao S, Chang EF, Teng CL, Heiser MA, Merzenich MM. Emergent categorical
1403 representation of natural, complex sounds resulting from the early post-natal sound
1404 environment. Neuroscience. 2013;248:30-42. Epub 20130606. doi:
1405 10.1016/j.neuroscience.2013.05.056. PubMed PMID: 23747304; PubMed Central PMCID:
1406 PMCPMC3838508.
1407 24. Bao S, Chang EF, Woods J, Merzenich MM. Temporal plasticity in the primary
1408 auditory cortex induced by operant perceptual learning. Nat Neurosci. 2004;7(9):974-81.
1409 Epub 20040801. doi: 10.1038/nn1293. PubMed PMID: 15286790.
1410 25. Han YK, Kover H, Insanally MN, Semerdjian JH, Bao S. Early experience impairs
1411 perceptual discrimination. Nat Neurosci. 2007;10(9):1191-7. Epub 20070729. doi:
1412 10.1038/nn1941. PubMed PMID: 17660815.
1413 26. Anbuhl KL, Yao JD, Hotz RA, Mowery TM, Sanes DH. Auditory processing remains
1414 sensitive to environmental experience during adolescence in a rodent model. Nat Commun.
1415 2022;13(1):2872. Epub 20220524. doi: 10.1038/s41467-022-30455-9. PubMed PMID:
1416 35610222; PubMed Central PMCID: PMCPMC9130260.
1417 27. Jaramillo S, Zador AM. The auditory cortex mediates the perceptual effects of
1418 acoustic temporal expectation. Nat Neurosci. 2011;14(2):246-51. Epub 20101219. doi:
1419 10.1038/nn.2688. PubMed PMID: 21170056; PubMed Central PMCID: PMCPMC3152437.
1420 28. Kato HK, Gillet SN, Isaacson JS. Flexible Sensory Representations in Auditory Cortex
1421 Driven by Behavioral Relevance. Neuron. 2015;88(5):1027-39. Epub 20151112. doi:
1422 10.1016/j.neuron.2015.10.024. PubMed PMID: 26586181; PubMed Central PMCID:
1423 PMCPMC4670799.
1424 29. Rothschild G, Nelken I, Mizrahi A. Functional organization and population dynamics
1425 in the mouse primary auditory cortex. Nat Neurosci. 2010;13(3):353-60. Epub 20100131. doi:
1426 10.1038/nn.2484. PubMed PMID: 20118927.
1427 30. Cohen L, Mizrahi A. Plasticity during motherhood: changes in excitatory and
1428 inhibitory layer 2/3 neurons in auditory cortex. J Neurosci. 2015;35(4):1806-15. doi:
1429 10.1523/JNEUROSCI.1786-14.2015. PubMed PMID: 25632153; PubMed Central PMCID:
1430 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.

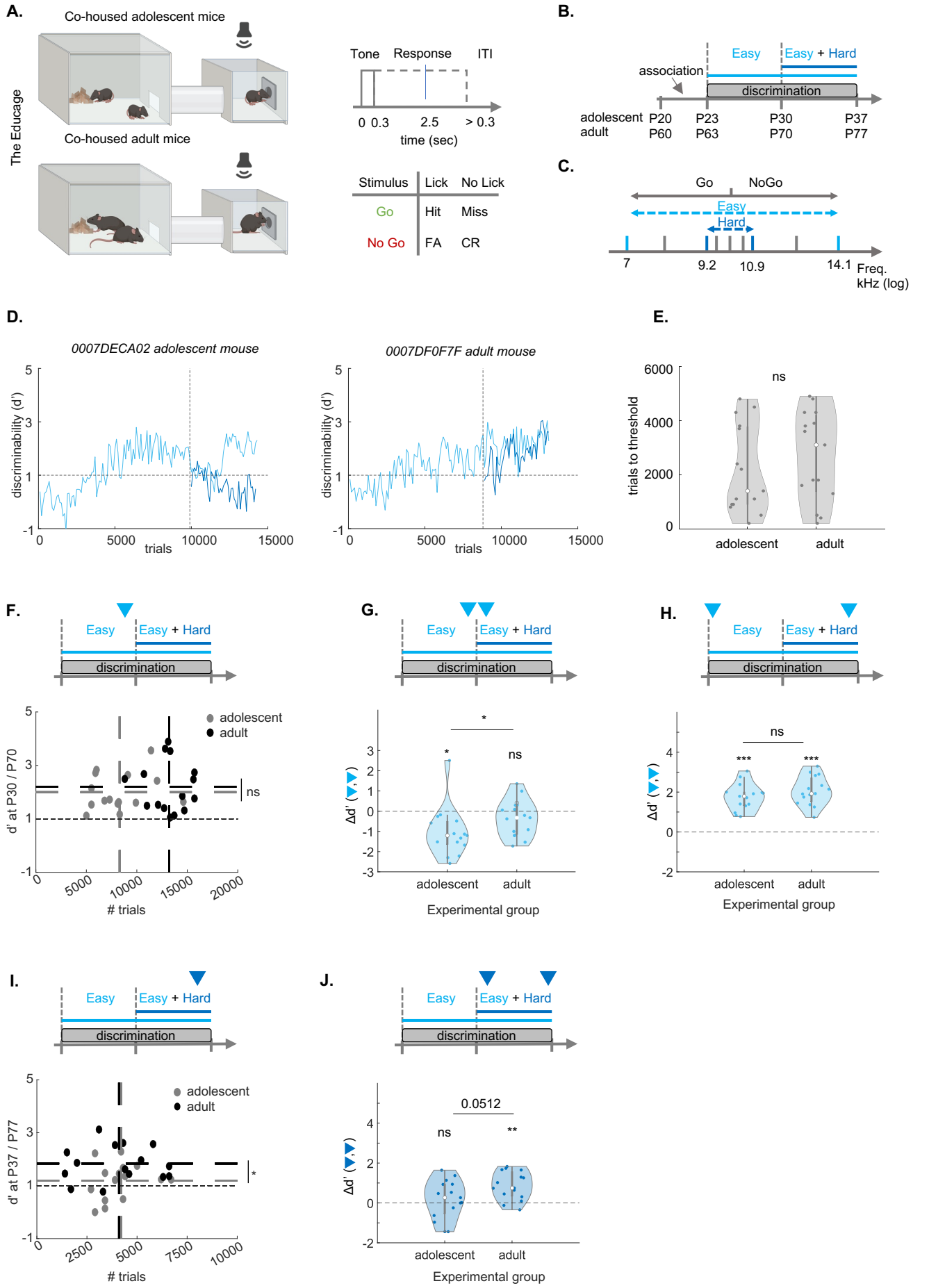
A.
Co-housed adolescent mice
The Educage
Co-housed adult mice
Tone
Response
ITI
0 0.3
2.5
> 0.3
time (sec)
Stimulus
Lick
No Lick
Go
Hit
Miss
No Go
FA
CR
B.
Easy
Easy + Hard
association
discrimination
adolescent
adult
P20
P60
P23
P63
P30
P70
P37
P77
C.
Go
NoGo
Easy
Hard
7
9.2
10.9
14.1
Freq.
kHz (log)
D.
0007DECA02 adolescent mouse
0007DF0F7F adult mouse
discriminability (d')
trials
E.
trials to threshold
ns
adolescent
adult
F.
Easy
Easy + Hard
discrimination
adolescent
adult
d' at P30 / P70
ns
trials
G.
Δd' (▼,▼)
*
ns
Experimental group
H.
***
I.
d' at P37 / P77
*
J.
0.0512
**

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; n trials=8317±712) and adult mice at P70 (black; n=15, d'=2.1986 ± 0.2441; n 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, vertical line), and mean d' per age group (z=-0.2074; p=0.8357, two-sample Wilcoxon rank sum test, 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; n trials=4163 ± 297; adult-black; n=15, d'=1.8342 ± 0.1743; n trials=4102±475) (mean trials per group: t-stat=0.1306, df=28, p=0.8970, two-sample independent t-test, vertical line; mean d' per group: z=-2.2398; p=0.0251, two-sample Wilcoxon rank sum test, 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	SEM	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).

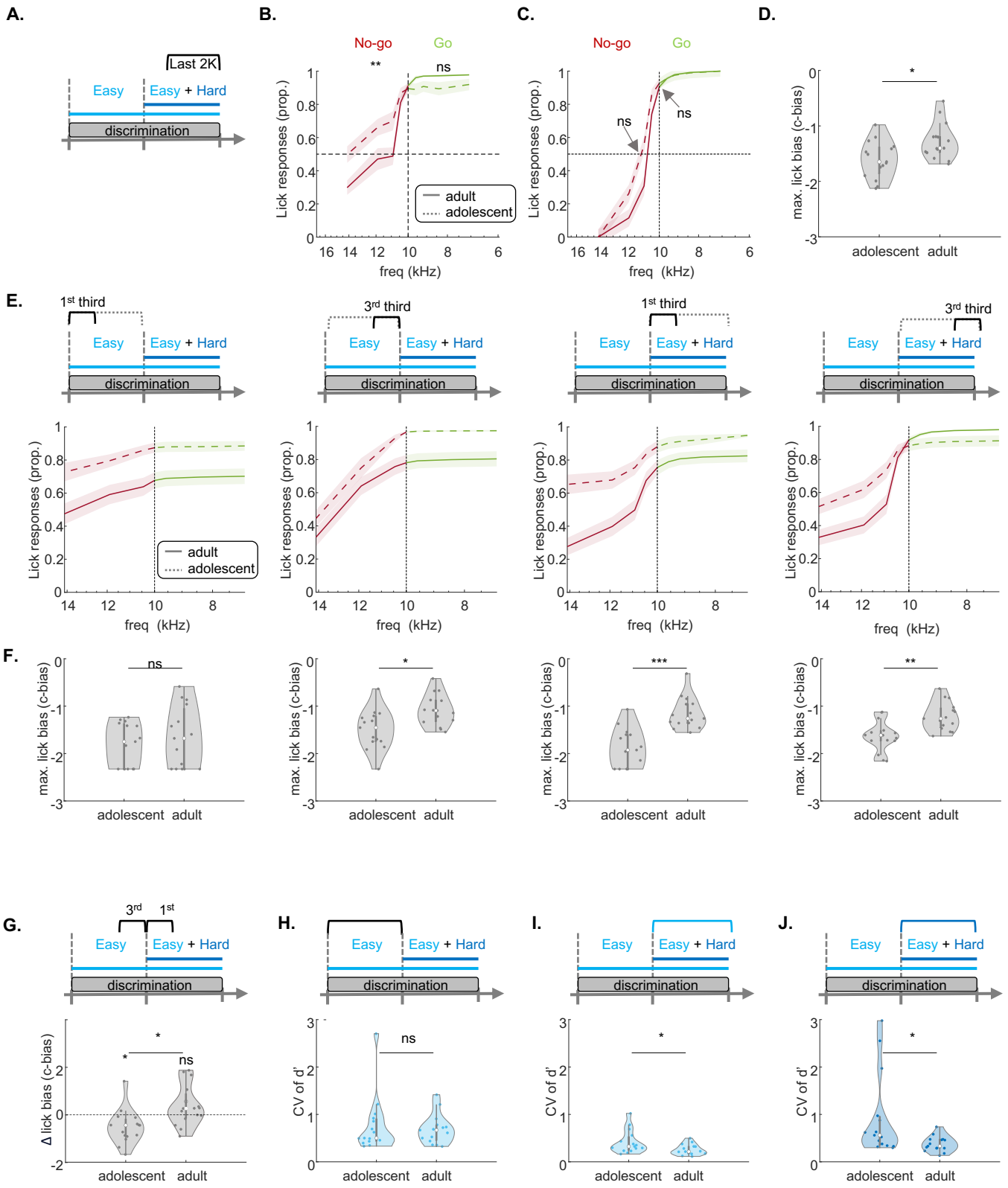
A.
Last 2K
Easy
Easy + Hard
discrimination
B.
No-go
Go
**
ns
Lick responses (prop.)
adult
adolescent
freq (kHz)
C.
No-go
Go
ns
ns
Lick responses (prop.)
freq (kHz)
D.
*
max. lick bias (c-bias)
adolescent
adult
E.
1st third
3rd third
1st third
3rd third
Easy
Easy + Hard
discrimination
Lick responses (prop.)
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 function0(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).

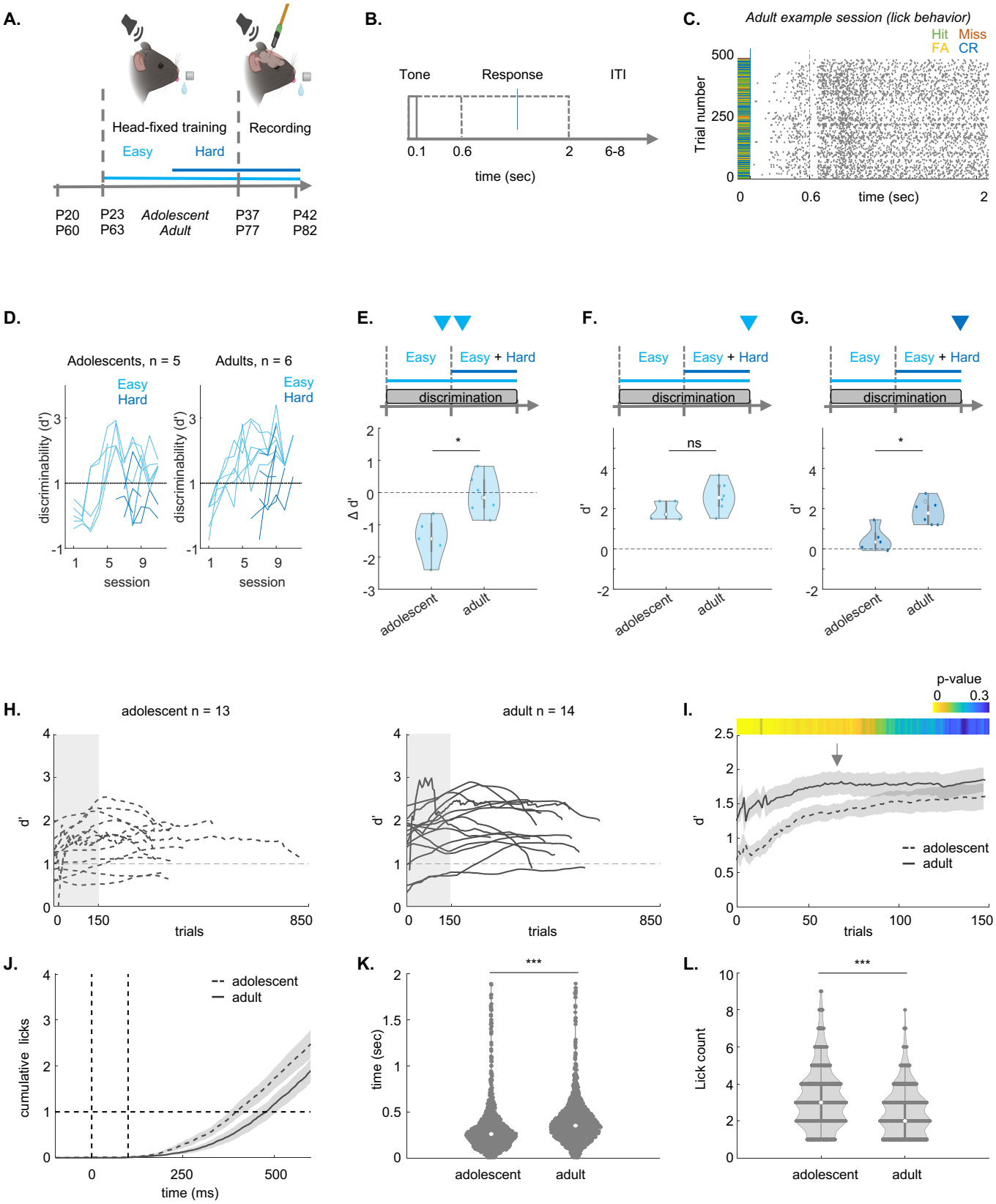
A.
Head-fixed training
Recording
Easy
Hard
P20
P60
P23
P63
Adolescent
Adult
P37
P77
P42
P82
B.
Tone
Response
ITI
0.1
0.6
2
6-8
time (sec)
C.
Adult example session (lick behavior)
Hit
Miss
FA
CR
Trial number
500
250
0
0
0.6
time (sec)
2
D.
Adolescents, n = 5
Adults, n = 6
Easy
Hard
discriminability (d')
session
E.
Easy
Easy + Hard
discrimination
Δ d'
*
adolescent
adult
F.
ns
d'
G.
*
H.
adolescent n = 13
adult n = 14
d'
trials
I.
p-value
0
0.3
adolescent
adult
J.
cumulative licks
time (ms)
K.
***
time (sec)
L.
***
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; 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.

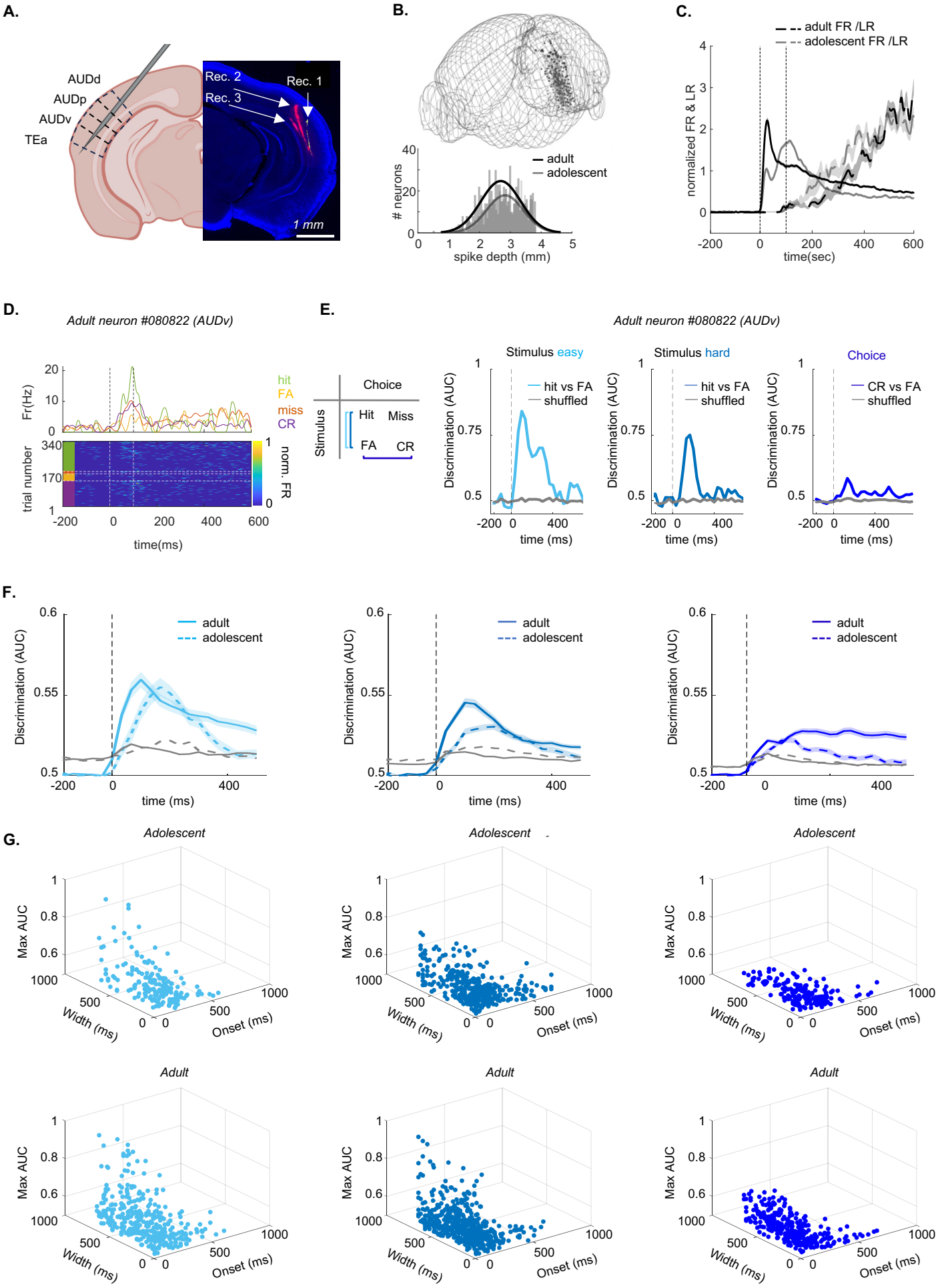
A.
AUDd
AUDp
AUDv
TEa
Rec. 2
Rec. 1
Rec. 3
1 mm
B.
neurons
adult
adolescent
spike depth (mm)
C.
adult FR /LR
adolescent FR /LR
normalized FR & LR
time(sec)
D.
Adult neuron #080822 (AUDv)
Fr(Hz)
hit
FA
miss
CR
trial number
norm. FR
time(ms)
E.
Adult neuron #080822 (AUDv)
Choice
Stimulus
Hit
Miss
FA
CR
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
Width (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; light green) and adults (n = 14; dark green). 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 and lick-rate (LR) 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 pertrial 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 (+- SEM) 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	StandardError	tStatistic	DF	pValue	CI lower	CI upper
Intercept	163.6104	7.9457	20.591	522	6.29587E-69	148.0009	179.2199
Onset (ms)	-29.7978	9.7626	-3.0523	522	0.00716101	-48.9765	-10.6191
Intercept	216.2921	13.2858	16.2799	522	5.15393E-48	190.1919	242.3924
Duration (ms)	63.2599	16.3499	3.8691	522	0.000369146	31.1401	95.3796
Intercept	0.6035	0.007	85.6404	522	0.0001	0.5897	0.6174
Max AUC	0.0172	0.0087	1.9776	522	0.145507056	0.0001	0.0342
Stimulus hard							
Fixed Effects	Estimate	StandardError	tStatistic	DF	pValue	CI lower	CI upper
Intercept	151.1905	5.4684	27.6479	941	3.6685E-123	140.4588	161.9222
Onset (ms)	-35.3357	7.1998	-4.9079	941	3.25392E-06	-49.4651	-21.2063
Intercept	220.614	8.4357	26.1525	941	2.8344E-113	204.0591	237.1689
Duration (ms)	55.6268	11.1065	5.0085	941	1.9647E-06	33.8305	77.4231
Intercept	0.5821	0.0034	170.2777	941	0.0001	0.5754	0.5888
Max AUC	0.0098	0.0045	2.1862	941	0.087143965	0.001	0.0187
Choice							
Fixed Effects	Estimate	StandardError	tStatistic	DF	pValue	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	0.0166	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 lantecy (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).

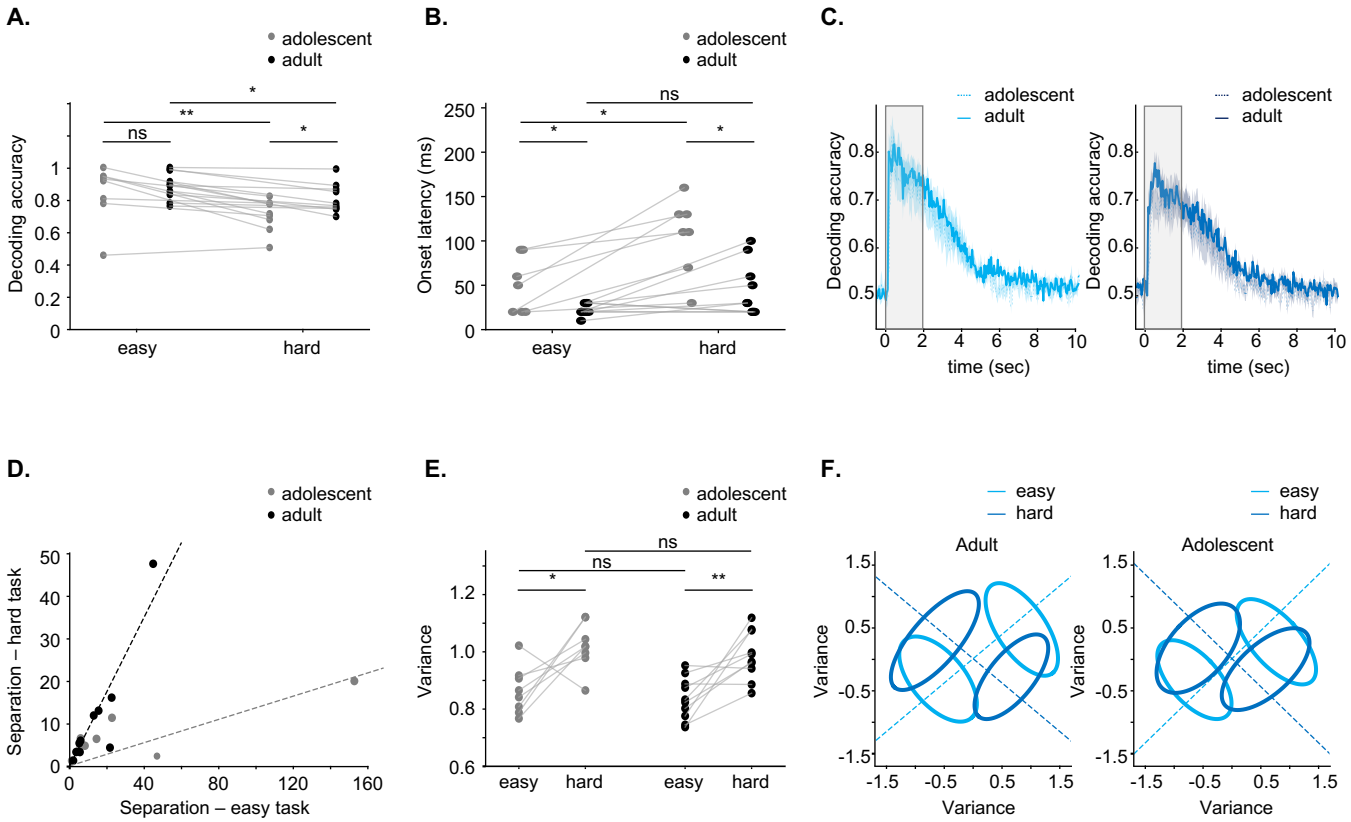

Figure 5: *Decoding in adult neuronal populations outperforms decoding in adolescents.* **A.** Decoding accuracy for the first 200ms across all recordings in both adults (black) and adolescents (grey) for the easy task (adolescents compared to adults, p=0.5000, Student's t-test) and the hard task (adolescents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

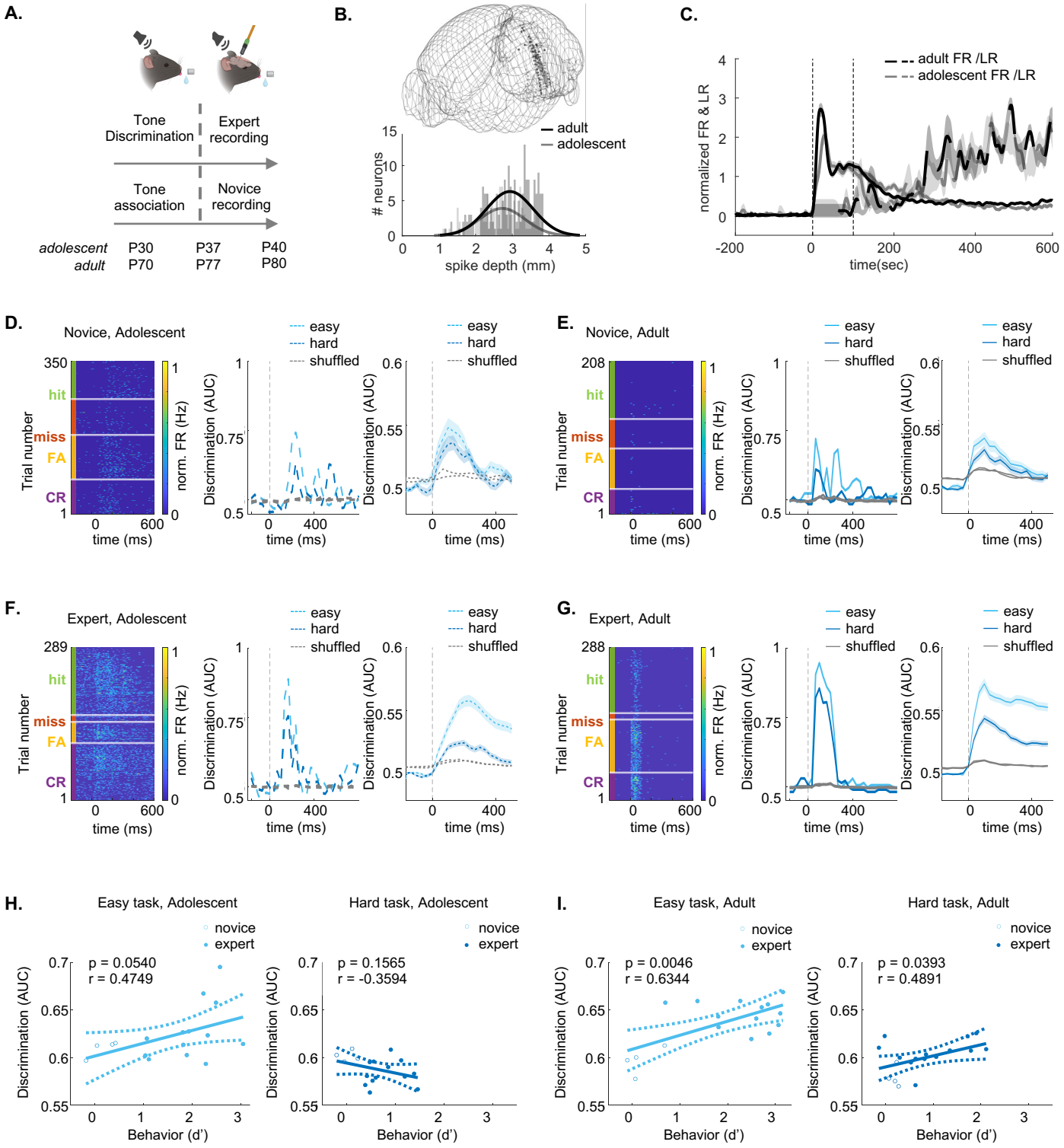

Figure 6: *Cortical activity during behavior reflects both age- and learning-induced effects.* **A.** Training and recording schedule for novice mice, compared to expert mice. **B.** 3D-Reconstruction of recording sites in novice adolescent (n = 6; grey) and novice adult (n = 6; black) mice. Bottom: spike-depth of excitatory tone-responsive L5/6 adolescent (n = 107; grey) and adult (n = 177; 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.

Fixed Effects Onset Latency(ms)	Estimate	SEM	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**	**SEM**	**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**	**SEM**	**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).

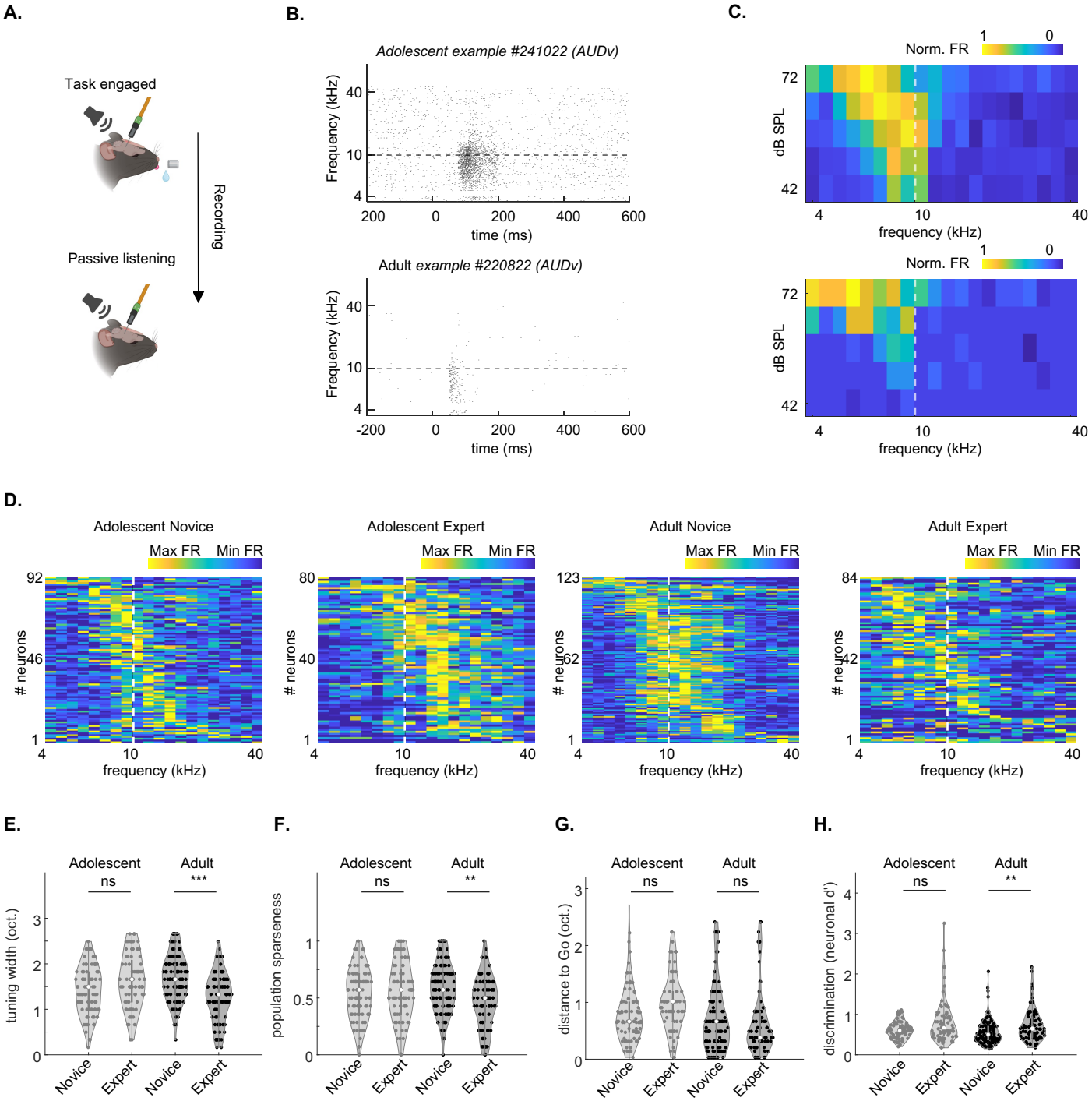

Figure 7: *Unlike adults, neuronal tuning properties of adolescents do not change after learning* ***A.*** Schematic showing that for the passive listening protocol, we continued our recording following the session of the engaged task (i.e. in satiated mice) by removing the waterspout. **B.** Example 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).