

**1 TITLE:**

**2**

**3 Distinct spatially organized striatum-wide acetylcholine dynamics for the learning and  
4 extinction of Pavlovian associations**

**5**

**6 AUTHORS**

**7** Safa Bouabid<sup>1</sup>, Liangzhu Zhang<sup>1</sup>, Mai-Anh T. Vu<sup>1</sup>, Kylie Tang<sup>1</sup>, Benjamin M. Graham<sup>1</sup>, Christian  
**8** A. Noggle<sup>1</sup>, Mark W. Howe<sup>1,\*</sup>

**9**

**10 AFFILIATIONS**

**11** <sup>1</sup>Department of Psychological & Brain Sciences, Boston University, Boston, MA, USA

**12** \*Lead contact; Correspondence: [mwhowe@bu.edu](mailto:mwhowe@bu.edu)

**13**

**14 ABSTRACT**

**15**

**16** Striatal acetylcholine (ACh) signaling has been proposed to counteract reinforcement signals to  
**17** promote extinction and behavioral flexibility. ACh dips to cues and rewards may open a temporal  
**18** window for associative plasticity to occur, while elevations may promote extinction. Changes in  
**19** multi-phasic striatal ACh signals have been widely reported during learning, but how and where  
**20** signals are distributed to enable region-specific plasticity for the learning and degradation of  
**21** cue-reward associations is poorly understood. We used array fiber photometry in mice to  
**22** investigate how ACh release across the striatum evolves during learning and extinction of  
**23** Pavlovian associations. We report a topographic organization of opposing changes in ACh  
**24** release to cues, rewards, and consummatory actions across distinct striatum regions. We  
**25** localized reward prediction error encoding in particular phases of the ACh dynamics to a  
**26** specific region of the anterior dorsal striatum (aDS). Positive prediction errors in the aDS were  
**27** expressed in ACh dips, and negative prediction errors in long latency ACh elevations. Silencing  
**28** aDS ACh release impaired behavioral extinction, suggesting a role for ACh elevations in  
**29** down-regulating cue-reward associations. Dopamine release in aDS dipped for cues during  
**30** extinction, but glutamate input onto cholinergic interneurons did not change, suggesting an  
**31** intrastriatal mechanism for the emergence of ACh elevations. Our large scale measurements  
**32** indicate how and where ACh dynamics can shape region-specific plasticity to gate learning and  
**33** promote extinction of Pavlovian associations.

**34**

**35 KEYWORDS**

**36** acetylcholine, dopamine, glutamate, striatum, neuromodulator dynamics, multi-fiber photometry,  
**37** behaving mice, pavlovian learning, behavioral flexibility, reward prediction error

38 INTRODUCTION

39

40 Cholinergic interneurons (CINs) supply the main source of acetylcholine (ACh) to the striatum,  
41 the principal input nucleus of the basal ganglia<sup>1,2</sup>. ACh release is capable of modulating the  
42 excitability and plasticity of striatal spiny projection neurons to influence behavior on timescales  
43 ranging from 100s of ms to days<sup>3,4</sup>. Electrophysiological studies have identified putative CINs by  
44 their characteristic tonically active firing profiles to investigate their natural dynamics in behaving  
45 rodents and primates during associative learning<sup>5,6</sup>. These studies revealed multi-phasic  
46 response profiles in CINs to primary rewards and reward predicting cues consisting of a short  
47 latency burst in firing, followed by a pause, then a post-inhibitory rebound<sup>7-9</sup>. Each of these  
48 components are variably expressed across CINs and can evolve in magnitude and duration as  
49 cues become associated with reward<sup>8,10,11</sup>. In particular, pauses in CIN firing to reward-predictive  
50 cues become larger and broader from the naive to learned stages of Pavlovian conditioning<sup>8,11</sup>.  
51 Larger cue-evoked CIN pauses for reward-associated cues are inverse in polarity to the  
52 coincident phasic elevations in cue-evoked dopamine (DA) neuron firing and release, widely  
53 hypothesized to represent positive reward prediction errors (RPEs)<sup>12-18</sup>. Positive RPEs encoded  
54 in DA release are believed to promote striatal synaptic plasticity for associative learning, while  
55 coincident dips in ACh release are believed to set a permissive temporal window for DA driven  
56 plasticity to occur<sup>19-21</sup>. Evidence in slices and *in-vivo* suggest that DA release may itself promote  
57 larger ACh pauses through inhibitory D2 receptors on CINs<sup>8,13,14</sup>. However, studies using  
58 different associative learning paradigms have found little evidence that CIN (or ACh) pauses are  
59 fully consistent with RPEs for both predictive cues and rewards, indicating that DA elevations  
60 and ACh pauses are at least partially independent signals<sup>10,15,22</sup>.

61

62 While *in-vivo* recordings of striatal ACh signaling have focused primarily on dynamics during  
63 initial associative learning, selective CIN manipulations in rodents have indicated roles in  
64 behavioral flexibility, response suppression, and extinction<sup>23-33</sup>. The specific effects of such  
65 manipulations vary across functionally distinct striatum regions. A key element of flexible  
66 behavior is the ability to down-regulate learned associations between previously valued cues  
67 and rewards when the associations are no longer valid, such as when a reward source is  
68 depleted. Dips in striatal DA release, putatively encoding negative prediction errors, arise to  
69 unexpectedly omitted rewards and to low predictive value cues and may promote the extinction  
70 of cue-reward associations<sup>34-37</sup>. However, although CINs have been implicated broadly in  
71 flexibility and extinction learning, few studies have attempted to examine the natural ACh  
72 dynamics as cues lose their reward predictive value.

73

74 One major challenge in understanding the dynamics of ACh signaling *in-vivo* is capturing  
75 variations in signaling across striatum regions. The striatum is a large, deep brain structure, and  
76 different functions in learning and action have been linked to particular striatal territories.  
77 Functional heterogeneity likely arises, at least in part, from topographically organized input from  
78 cortical and sub-cortical regions. Mounting evidence has indicated that anatomical  
79 heterogeneity is also manifested in striatal ACh signaling during behavior. Although some  
80 studies have shown periods of high synchrony across CINs within a given region<sup>6,8,13</sup>, others  
81 have reported significant signal differences across striatum regions<sup>22,38-43</sup>. Anatomical

82 heterogeneities in ACh signaling have been difficult to assess due to the sparseness of CINs in  
83 the striatum (~3-5% of total population) and inherent limitations in spatial coverage of current  
84 optical and electrophysiological approaches. Standard electrode implants measure spiking from  
85 only a few unique CINs per subject, typically target limited striatal regions (< 4), and without  
86 optical tagging can suffer from ambiguities in accurate CIN identification. Fiber photometry has  
87 been used to optically measure ACh release, but similarly has been limited to 1-2 striatal  
88 regions per experiment. Therefore, it is poorly understood how and whether ACh release  
89 dynamics differ across the striatum volume to support distinct functions in associative learning  
90 and extinction.

91

92 To address these limitations, we applied a multi-fiber array technique to optically measure rapid  
93 ACh release dynamics across over 50 locations simultaneously throughout the entire  
94 3-dimensional striatum volume in head-fixed, behaving mice. Changes in multi-phasic ACh  
95 signals over learning and extinction were cue modality and signal component specific and  
96 varied across distinct anatomical gradients. Signaling at specific locations in the anterior dorsal  
97 striatum (aDS) was consistent with positive and negative RPEs for particular ACh signal phases  
98 and was opposite in polarity to classic DA RPEs. In support of a functional role for elevations in  
99 aDS ACh release in the down-regulation of cue-reward associations, silencing ACh release in  
100 the aDS impaired extinction but left initial learning intact. Finally, measurements of aDS DA and  
101 glutamate release onto CINs suggest an intrastriatal mechanism for the emergence of aDS ACh  
102 elevations. Overall, our results identify a spatiotemporal organization of striatal ACh signals  
103 positioned to gate region-specific plasticity during learning and downregulate learned  
104 cue-reward associations during extinction.

105

## 106 RESULTS

107

### 108 *Striatum-wide measurements of ACh release with chronic micro-fiber arrays*

109

110 We optically measured rapid ACh release dynamics across the striatum in head-fixed behaving  
111 mice using large-scale arrays of 55-99 optical micro-fibers<sup>44</sup> (50µm diameter) distributed  
112 uniformly across the 3-dimensional striatum volume (Fig. 1a-e). Each fiber collected  
113 fluorescence from post-synaptic neurons expressing the genetically encoded ACh sensor,  
114 GRAB-ACh3.0<sup>45</sup> (Fig. 1a). Based on previous empirical measurements and light scattering  
115 models<sup>46,47</sup>, we estimate that our fibers collect fluorescence from a tapered collection volume  
116 extending approximately 100µm axially and 25µm radially from each fiber tip (see Vu et al, 2024  
117 for details<sup>48</sup>). The minimum separation of fiber tips in our arrays was 220µm radially and 250µm  
118 axially ensuring no (or very minimal) overlap in the collection volumes for each fiber. Fiber  
119 locations in the brain were precisely reconstructed from post-mortem computerized tomography  
120 scans and aligned to a common coordinate framework to enable comparisons of spatial patterns  
121 of ACh release within and across subjects (Fig. 1d,e)<sup>48,49</sup>. Rapid, bi-directional changes in  
122 fluorescence, a proxy for changes in ACh release and synchronous firing of cholinergic  
123 interneurons<sup>50</sup>, could be reliably detected across striatum locations (Fig. 1c). Unpredicted water  
124 reward deliveries produced multi-phasic release profiles at some locations, consisting of an  
125 early peak, followed by a dip below baseline, then a later peak (Figs. 1f, g). However, all three

126 signal components were not present in every region (Fig. 1g). The magnitudes of each  
127 component were generally stable across days enabling us to account for potential changes in  
128 sensor expression or light collection during chronic recordings. Only a small fraction of fibers  
129 showed statistically significant changes in unpredicted reward mean  $\Delta F/F$  across days of  
130 learning (1.6%, 5.7% and 5.7% of the 295 total fibers within the striatum for early peak, dip and  
131 late peak, respectively,  $p < 0.01$ , one-way ANOVA test comparing the mean of each day with the  
132 mean of every other day, see Methods, Fig. 1f). Thus, our optical approach enabled us to  
133 investigate the evolution of simultaneous, spatiotemporal patterns of ACh release across the  
134 striatum during learning.

135

136 To control for potential signal contamination by hemodynamic or motion artifacts, we conducted  
137 measurements with quasi-simultaneous 405nm illumination, the approximate isosbestic point of  
138 the sensor<sup>45</sup> and in mice expressing a non-functional null version of the ACh3.0 sensor (Fig. 1h,  
139 i). Rapid peaks and dips were generally not observed in these control conditions. However,  
140 some slow, low amplitude fluctuations were present, which could largely be eliminated with high  
141 pass filtering (see Methods, Fig. 1i). In mice expressing the null sensor, we found that some  
142 small artifacts were present with 470nm illumination which were not detected with 405nm  
143 illumination, and overall, the artifactual fluorescence changes measured with 405nm illumination  
144 were only weakly correlated with 470nm illumination (Fig. 1h). Therefore, quasi-simultaneous  
145 405nm illumination may not reliably account for all potential signal artifacts, likely because of  
146 differences in hemoglobin absorbance and tissue scattering between the two illumination  
147 wavelengths<sup>51</sup>. Thus, going forward, we used the null sensor expressing cohort with 470nm  
148 illumination to rule out significant contributions of artifacts to behavior related signals in ACh3.0  
149 expressing mice.

150

### 151 *Spatially organized, opposing changes in striatal ACh release during Pavlovian learning*

152

153 To determine how and where learning related changes in cue evoked ACh release dynamics  
154 occur, we mapped changes in ACh release across the striatum as mice learned a dual-cue  
155 delay Pavlovian conditioning task. Head-fixed mice ( $n = 8$ ) on a spherical treadmill<sup>52</sup> were  
156 trained to associate light (470nm LED, 7mW) and tone (12kHz, 80dB) stimuli with water reward  
157 after a fixed 3s delay (Fig. 2a). Cues were presented one at a time and pseudo randomly  
158 interleaved with 4-40s intertrial intervals (ITIs) drawn from a uniform distribution. Non-contingent  
159 rewards were delivered occasionally (8 per session) in ITI periods. During Pavlovian learning,  
160 mice optimized their licking in two ways: they increased the fraction of time spent licking during  
161 the cue period ( $p = 0.0078$ , two-tailed Wilcoxon signed rank test, Fig. 2b-d), and they decreased  
162 non-contingent, ‘spontaneous’ licking during the ITI period ( $p = 0.0273$ , two-tailed Wilcoxon  
163 signed rank test, Fig. 2e). A lick index was computed, reflecting the relative proportion of  
164 anticipatory licking within the cue period relative to the ITI, to assess learning of the cue-reward  
165 association on each session (Fig. 2c).

166

167 Consistent with previous findings, we observed, for some fibers, multi-phasic, positive and  
168 negative changes in ACh release to conditioned cues after learning, consisting of a fast latency  
169 ('early') peak, a dip, and a second longer latency ('late') peak following the dip (Fig. 2f, h;

170 Fig.S1a, b). Prior to learning, cues evoked multi-phasic ACh release at many fewer fibers for all  
171 three components (Fig. 2f-i; Fig.S1a-d ). Early ACh peaks emerged and increased in magnitude  
172 with learning for a large fraction of fibers for both cues (light: 61.69%, tone: 51.53% of total  
173 fibers), while a smaller fraction exhibited peaks that remained stable across learning (Fig. 2f,g;  
174 Figs. S1c,d and S2c). In contrast, ACh dips for both cues predominantly became larger with  
175 learning (Fig. 2h,i; Fig.S2c, light: 46.78%, tone: 39.66% of total fibers). Late ACh peaks also  
176 emerged, but for a much larger fraction of locations for the light cue than the tone (Fig.S2a,c;  
177 light: 37.63%, tone: 8.47% of total fibers). These learning related changes were consistently  
178 present across mice expressing the ACh3.0 sensor but were not observed in mice expressing  
179 the non-functional mutant sensor, ruling out significant contributions of hemodynamic or motion  
180 artifacts (Fig. S2j).

181

182 To visualize anatomical patterns in the evolution of cue-evoked ACh release across the striatum,  
183 we constructed maps for each component reflecting spatially weighted averages of release  
184 changes during learning across fibers (see Methods). Signal changes for the early peak and dip  
185 were concentrated across two spatially segregated, partially overlapping regions (Fig. 2j,k; Fig.  
186 S2d). For both cues, the largest increases in the early peak were concentrated in a region of the  
187 ventral striatum (VS), near the center of the striatum in the A/P axis (Fig. 2j; Figs. S1e,g and  
188 S2d-f). While some early peak changes were present in the anterior dorsal striatum (aDS), the  
189 changes were smaller and peaks had shorter latencies compared to the VS (Fig. 2j; Figs. S1e-g  
190 and S2d-f). In contrast to the early peak, the largest changes in the dip were concentrated in the  
191 aDS for the light but shifted slightly posterior for the tone (Fig.2k; Fig. S2d-f). Dip changes also  
192 extended partially into the most anterior portion of the VS (Fig. S1g and S2e,f). Increases in the  
193 late peak overlapped with dip changes for the light cue in the aDS, but the sparse late peaks for  
194 the tone cue were confined to a restricted region of the posterior striatum (Fig. S2b,d-f).  
195 Variation in the component changes across striatal axes was quantified using a generalized  
196 linear model which included a mouse identity term to account for differences across individual  
197 mice (Fig. 2l,m, see Methods). Spatial gradients for early peaks and dips differed significantly  
198 across distinct striatal axes (Fig. 2l). In general, more significant components and larger  
199 changes with learning were present for the light cue relative to the tone (Fig. S1a-d and  
200 S2c,f,g,h). Additionally, gradients for tone and light cues differed significantly along at least one  
201 axis for all three signal components, and cue selective changes were present for individual  
202 fibers in distinct territories perhaps reflecting the organization of excitatory input to the  
203 striatum<sup>53,54</sup> (Fig. 2m; Fig. S2g). In summary, the acquisition of Pavlovian associations is  
204 accompanied by bi-directional (larger peaks, larger dips), sensory modality specific changes in  
205 cue-evoked ACh release across distinct 3-dimensional striatal territories.

206

207 Changes in cue-evoked ACh release with learning are partly consistent, in some regions, with  
208 bi-directional encoding of positive RPEs, as widely reported in DA neuron firing and release  
209 <sup>55-57</sup>. We next asked whether signal changes at reward delivery reflected positive RPE encoding.  
210 Like the cues, water delivery was associated with sequential peaks and dips in ACh release  
211 (Fig. 3a,b; Fig. S3a, b). However, unlike early ACh peaks at the cue which almost exclusively  
212 increased with learning at a majority of sites, peaks at reward changed for a minority and  
213 displayed a mixture of increases and decreases (Fig. 3b). In contrast, reward-evoked ACh dips

214 were significantly smaller for a majority of fibers (57.29%, Fig. 3a,b). Consistent with putative  
215 RPE encoding, the dip changes at reward were opposite to the changes observed for the cues  
216 (Fig. 2). Late peaks predominantly increased (inconsistent with RPE encoding), but weakly (Fig.  
217 3b). The changes in reward dips from pre to post learning were consistent with comparisons  
218 between unpredicted (random) and predicted (cue associated) rewards, and for rewards  
219 following low vs high probability cues during extinction (Fig. S3c,d). Diminished reward dips  
220 were observed over a broad region of the dorsal striatum extending across the regions showing  
221 cue-evoked dip changes for the light and tone cues (Fig. 3c; Fig.S3a,b). Increases and  
222 decreases to the peaks were scattered, with few concentrated in the VS regions showing strong  
223 elevations in the cue peak (Fig. 3c). Overall, only a relatively small proportion of fibers (light:  
224 13.9%, tone: 9.8% of total fibers) had both increases in ACh early peaks to the cue and  
225 decreases in peaks to the reward (indicative of positive RPE encoding), with the majority  
226 increasing exclusively for cues (Fig. 3d,e; Fig.S3d). For the dip component, a larger fraction of  
227 fibers (light: 29.2%, tone: 28.1% of total fibers) changed for both the cue and reward (smaller  
228 dip at reward, larger at cue) relative to cue or reward only, consistent with positive RPE  
229 encoding (Fig. 3d,e; Fig.S3e). Fibers with putative RPE encoding for the dip were localized to  
230 the aDS for the light cue and scattered across the more posterior dorsal striatum for the tone  
231 (Fig. 3d,e; Fig.S3e). Significant spatially organized changes were largely absent in null-sensor  
232 mice (Fig. S3f). In summary, our results indicate that putative positive RPEs to cues and  
233 rewards are encoded predominantly by ACh dips within restricted, sensory modality specific  
234 regions of the dorsal striatum.

235

### 236 *Emergence of spatially organized ACh signals during extinction*

237

238 We next tested how and where cue-evoked ACh release changes as cues lose their predictive  
239 value during partial extinction learning. Following initial Pavlovian conditioning, the reward  
240 probability associated with one of the two cues was downshifted to 20% and the other to 80%  
241 (Fig. 4a). Cue-evoked anticipatory licking decreased over days specifically for the lower  
242 probability cue, indicating successful cue-specific extinction learning, at which point, reward  
243 probabilities for each cue were reversed (Fig. 4b-d). ACh early peaks predominantly decreased  
244 relative to the post learning phase for the low probability cue (light: 36.27%, tone: 40.34% of  
245 total fibers, Fig.4e,f; Fig S4c) and dips became smaller (light: 22.03%, tone: 29.5% of total  
246 fibers, Fig. 4g,h; Fig S4c). The direction of these changes represented a reversal of the changes  
247 during initial learning (Fig. 2, Fig.S2). However, unlike the early peaks and dips, ACh late peaks  
248 primarily increased for the low probability cue relative to late learning for both cues (light:  
249 27.46%, tone: 20.68% of total fibers, Fig. 4g,h; Fig S4c). Differences were consistent for the  
250 initial extinction and reversal phases, so data was combined for all analyses. These results  
251 indicate that the degradation of cue-reward associations is accompanied by a reversal of  
252 learned dynamics for early peak and dip components but an emergence of new ACh increases  
253 for the late peak component.

254

255 We then examined the spatial organization of the changes in each component during extinction.  
256 The reversal of the early peak and dip changes were concentrated in largely the same regions  
257 of the VS and aDS (for the light cue) respectively as the regions showing changes during initial

258 learning (Figs. 2j,k and 4i,j; Fig. S4a-c). The increases in the late peaks were concentrated  
259 almost exclusively in a restricted region of the aDS (Fig. 4j, Fig. S4a-c). These regional patterns  
260 held for differences in signaling between the 20% and 80% probability phases for a given cue  
261 modality (Fig. S5). The spatial gradient of the late peak changes was significantly different from  
262 that of the early peak changes for both cue modalities (Fig. 4k). Moreover, the spatial  
263 distribution of changes differed significantly between tone and light cues for the dip and late  
264 peak but not for the early peak (Fig. 4l; Fig. S4e,f). The region-specific emergence of late peaks  
265 in aDS occurred many trials before the decreases in anticipatory spout licking to the 20%  
266 probability cue (light cue: median= 29 trials before licking decrease, STD=27, Fig. 4m,n; tone  
267 cue: median = 2 trials before licking decrease, STD =15, Fig. 4o,p), indicating that these signals  
268 are positioned to drive learning related plasticity, perhaps contributing to behavioral extinction  
269 (see Discussion). Again, significant changes during extinction were not present in null sensor  
270 expressing mice (Fig S4d). Overall, these findings indicate that flexible downshifting of learned  
271 cue values is accompanied by distinct, bi-directional changes in cue-evoked ACh release across  
272 different striatal regions that precede a downregulation of behavior. Most notably, aDS late  
273 peaks did not revert back to a pre-learning state (as for peaks in the VS) but acquired new  
274 elevations reflecting the relative (negative) change in predictive cue value.

275

276 Increases in ACh late peaks for the low value cues during extinction are consistent with the  
277 emergence of negative RPE encoding reflecting either the negative change in predictive cue  
278 value from the learned phase or the relative difference between the high and low value cues.  
279 The aDS region exhibiting the emergence of cue-evoked late peaks during extinction aligned  
280 with the region encoding positive RPEs in ACh dips to light cues and rewards during initial  
281 learning (Fig. 3d), suggesting that the aDS may preferentially encode both positive and negative  
282 RPEs through ACh dips and late peaks respectively (see Discussion). Our experimental design  
283 did not allow us to unambiguously resolve putative negative RPE signals for reward omissions  
284 during the extinction phase, since mice were not required to precisely estimate the timing of  
285 reward delivery, and anticipatory licking began prior to the time of expected reward. However,  
286 alignment to reward omission could be tested during ITI periods, when mice spontaneously  
287 initiated spout licks when water was not present. Mice initiated spontaneous spout licking bouts  
288 in the ITI less frequently after learning the cue-reward association (Fig. 2e). This indicated that  
289 they predict a higher probability of random reward in the ITI early prior to learning the task  
290 structure. Therefore, unrewarded, spontaneous spout licks in the ITI would result in a larger  
291 negative RPE following the lick onset early in learning. Lick bout initiations in the ITI pre learning  
292 were associated with multi-phasic changes in ACh release at distinct latencies (Fig. S6a,b).  
293 Short latency increases (early peaks) in ACh began before the tongue contacted the lick spout,  
294 around lick initiation (Fig.5a). Dips were present at intermediate latencies, while longer latency  
295 increases (late peaks) occurred several hundred milliseconds after spout contact (Fig. 5c). All  
296 three components changed with learning for a significant proportion of fibers. Early peaks and  
297 dips predominantly decreased (Fig. 5b, early peak: 40%, dip: 26.1% of total fibers), while late  
298 peaks showed a mixture of increases and decreases across fibers (Fig. 5d, 23.1% and 18%  
299 respectively). Early and late peaks were concentrated across largely distinct striatum locations  
300 prior to learning. Early peaks were prominent in the posterior striatum, while late peaks were  
301 localized primarily to the aDS (Fig. S6b) and peaks in both regions decreased with learning (Fig.

302 5e,f; Fig. S6c). Some late peaks also emerged and became larger with learning, but in a distinct  
303 region of the central lateral striatum (Fig. 5f; Fig.S6c). Dips became larger with learning across  
304 the dorsal striatum (Fig. 5e, Fig.S6c). The changes in the ACh late peaks aligned in the aDS  
305 with the late peak elevations to the downshifted cues during extinction (Fig. 5g). Together, these  
306 results indicate that ACh release in the aDS signals putative negative RPEs through elevations  
307 in long-latency ACh peaks and positive RPEs through dips (Fig. 5g,h).

308

309 *Changes in ACh release during learning and extinction cannot be explained by locomotion*  
310 *changes*

311

312 ACh signaling has been previously linked to changes in locomotion <sup>39,50,58,59</sup>, so we asked  
313 whether changes we observed across learning phases could be explained by changes in  
314 treadmill velocity. A generalized linear model (GLM) was used to partially account for potential  
315 correlations with continuous linear velocity and acceleration (see Methods), but we conducted  
316 additional analyses to address potential movement contributions not captured by the GLM.  
317 During initial Pavlovian learning, mice decelerated significantly more, on average, after cue  
318 onset and less after reward delivery as they learned the cue-reward associations (Fig. S7a,b,  $p$   
319  $> 0.05$  for peak deceleration and  $p < 0.05$  for velocity changes after cue onset, Friedman test;  $p$   
320  $< 0.05$  for peak deceleration and  $p > 0.05$  for velocity after reward delivery, Wilcoxon rank-sum  
321 test). Differences were not present over learning for spontaneous ITI lick bouts (Fig. S7c,  
322  $p > 0.05$ , Wilcoxon rank-sum test) or during extinction for high and low probability cues or relative  
323 to the late learning phase (Fig. S7a,  $p > 0.05$ , Friedman test). To address the possibility that  
324 locomotion deceleration signals may contribute to the region-specific ACh changes that we  
325 report with initial learning (Figs. 2 and 3; Figs. S1-3), we first examined ACh release to  
326 spontaneous decelerations in the ITI period with similar magnitude to decelerations at cues and  
327 rewards (Fig. S7d-g). At some locations, primarily in the dorsal striatum, we observed bi-phasic  
328 peaks and dips in ACh release associated with ITI decelerations (Fig. S7d). However, these  
329 peaks and dips in  $\Delta F/F$  were present across fewer fibers and were much lower in magnitude  
330 relative to changes to cues and rewards with learning (Fig. S7e-g). Most fibers with large cue  
331 and reward changes with learning did not have any significant ITI deceleration response  
332 (36.15% and 23.76% of fibers with significant peak and dip changes at cue had significant  
333 maxima and minima for ITI decelerations respectively; 50% and 25.43% for fibers with peaks  
334 and dips at reward respectively; Fig. S7f,g). In addition, change magnitudes at cue and reward  
335 with learning were not correlated with magnitudes of peaks and dips to spontaneous  
336 decelerations across fibers (Fig. S7l,m; Table 1). For most fibers, early peaks and dips for cues  
337 and rewards were highly consistent in timing and amplitude across trials within a given session,  
338 despite significant variability in the timing and magnitude decelerations at cue and reward (Figs.  
339 2, 3; Fig.S7h,i). Moreover, ACh release to cues, rewards, and ITI licks were mostly insensitive to  
340 trial-by-trial variations in the size of the associated deceleration (Fig. S7h-k). ACh peaks and  
341 dips for only a small fraction of fibers with significant response components were significantly  
342 correlated with trial by trial deceleration for cues (7.4% of 270, 7.1% of 224, and 1.1% of 180  
343 fibers for early peak, dip, and late peak of light cue respectively; 8.3% of 262, 5.1% of 176, and  
344 3.9% of 76 locations for early peak, dip, and late peak of tone cue respectively; Pearson's  
345 correlation,  $p < 0.01$ ) and rewards (19.1% of 287, 23.5% of 276, and 6.1% of 229 locations for

346 early peak, dip, and late peak respectively; Pearson's correlation,  $p<0.01$ ). Thus, although  
347 changes in locomotion modulate ACh release at some striatal locations, locomotion related  
348 signaling per se could not account for the patterns of cue and reward evoked release that we  
349 observed.

350 *Intact ACh release in the aDMS is required for extinction of Pavlovian associations*

351

352 Our ACh release measurements identified transient increases in cue-evoked ACh release in the  
353 aDS (late peaks), consistent with negative RPEs, which may contribute to downshifting the  
354 predictive value of cues and the expression of appetitive behaviors, such as licking, when  
355 reward probabilities change (Figs. 4 and 5). We tested whether suppressing ACh release  
356 altered spontaneous and cue evoked spout licking during Pavlovian learning and partial  
357 extinction by virally expressing Tetanus toxin light chain (TeIC) selectively in cholinergic  
358 interneurons in the aDS of ChAT-cre mice. Targeting was informed by the distribution of late  
359 peak changes for the light cue (Fig. 4j), so injections were shifted slightly to favor the dorsal  
360 medial striatum (aDMS, see Methods, Fig. 6a-c). TeIC expression eliminates synaptic vesicle  
361 exocytosis in cholinergic interneurons, and consequently blocks ACh release<sup>60</sup> (Fig. 6c). ACh  
362 suppression did not affect the mice ability to learn the cue-reward association, as both TeIC and  
363 control mice increased the frequency of licking in the cue period over learning (Two-way  
364 repeated measure ANOVA,  $p = 0.02$  for Telc and  $p = 0.0001$  for control, Fig. 6d). This is  
365 expected, given that the primary changes in ACh release in aDS during initial learning were in  
366 the dips, which may be permissive for learning<sup>19,61-63</sup> (Fig. 2i,k; Fig. S2c,d). During cue specific  
367 extinction, control mice exhibited a down-regulation of anticipatory licking relative to the post  
368 learning phase, as expressed by the significant decrease in the fraction of time spent licking  
369 during cue period ( $p = 0.04$ , Two-way repeated measures ANOVA, Fig. 6d,f). However, TeIC  
370 expressing mice did not show a significant down-regulation of cue period licking relative to the  
371 learned phase ( $p = 0.93$ , Two-way repeated measures ANOVA, Fig. 6d,f). TeIC mice also had  
372 slightly higher licking rates during the ITI and cue periods than the control mice at all phases of  
373 learning, perhaps indicating a general difference in suppression of spontaneous licking (Fig.  
374 6d,e). Although the manipulation was not temporally specific, these data indicate that intact ACh  
375 release in the aDMS contributes to the extinction of learned Pavlovian responses following a  
376 downshift in cue associated reward probabilities.

377

378 *Cue and unrewarded lick evoked dopamine dips opposite to ACh late peaks in aDS*

379

380 We next investigated potential contributors to the emergence of the downshifted cue and  
381 spontaneous lick related ACh increases in the aDS. Dips in DA neuron firing and release have  
382 been measured in response to low value cues and unpredicted reward omissions (i.e. negative  
383 prediction errors) and have been proposed to contribute to extinction learning<sup>34,35,55,64</sup>. If present,  
384 these DA dips may disinhibit CINs through reduced tonic activation of inhibitory D2  
385 receptors<sup>13,60,65,66</sup>. We tested whether DA signals in the aDS are consistent with a potential role  
386 in facilitating the observed elevations in ACh signaling. We optically measured DA release  
387 dynamics with dLight1.3b<sup>67</sup> in the same regions of the aDS where we observed the increases in  
388 ACh peaks to unrewarded licks and down-shifted cues (Figs. 4,5,7a). Consistent with prior work,  
389 DA release rapidly increased to conditioned cues following learning (Fig. 7b,c). Following partial

390 extinction (Fig. 4a), dips below baseline emerged after short latency peaks, selectively for the  
391 low probability cue (Fig. 7b-d). The average latency to the trough of DA dips (median to trough =  
392 0.63s) preceded the average latency of the ACh late peaks (median = 0.8s) recorded in the  
393 same aDS region in the ACh3.0 expressing group (Fig. 7e,  $p = 0.02$ , Two-tailed rank-sum test).  
394 Like the ACh late peaks, dips in cue-evoked aDS DA emerged within only a few trials after the  
395 probability shift, well before the change in cue-evoked licking (median = 27 trials prior to lick  
396 change, STD=23.4 for light, median = 22 trials, STD=42.8 for tone, Fig. 7f). Like the ACh3.0  
397 expressing mice, there were no differences in the dLight1.3b group in cue-related decelerations  
398 between the post learning and extinction phases, so DA dips could not be explained by  
399 locomotion changes (Fig. 7g). We then asked whether DA release for spontaneous ITI licks also  
400 exhibited learning related dips reflecting shifts in negative RPE signaling with learning. DA  
401 release in the aDS increased for spontaneous licks, rising just prior to spout contact (Fig. 7h,i).  
402 Early in Pavlovian learning, peaks were followed by dips below baseline, but these dips largely  
403 disappeared after learning ( $p = 2.6 \times 10^{-3}$ , two-tailed Wilcoxon rank-sum test), mirroring opposite  
404 changes in ACh release as ITI licks became less frequent (Fig. 7h-j). Like the cue responses,  
405 DA dips had a shorter average latency (median = 0.33s) relative to the ACh late peaks (median  
406 = 0.44s,  $p = 3.33 \times 10^{-6}$ , two-tailed Wilcoxon rank-sum test, Fig. 7k). These results indicate that  
407 aDS DA release reflects putative negative prediction errors in dips to extinction cues and  
408 unrewarded ITI licks, which are opposite in polarity and precede the average latency of ACh late  
409 peaks in the same region.

410

411 *Changes in glutamate release onto aDS CINs cannot solely account for ACh increases during*  
412 *extinction*

413

414 Next, we asked whether emerging ACh increases to cues after partial extinction may develop as  
415 a consequence of increases in excitatory glutamatergic drive. To test this possibility, we  
416 expressed the genetically encoded glutamate sensor iGluSnFr<sup>68</sup> selectively in aDS cholinergic  
417 interneurons of ChAT-cre mice to measure rapid changes in glutamate release (GluCIN) during  
418 partial extinction (Fig. 8a-c). Interestingly, GluCIN release in the aDS to the light stimulus  
419 following Pavlovian learning was bi-phasic, consisting of an initial fast latency increase followed  
420 by a slower latency increase (Fig. 8d,e). Consistent with the relative absence of strong ACh  
421 release peaks in the aDS to the conditioned tone (Figs. S1d, S2d,f), no rapid changes in aDS  
422 GluCIN release were present for the tone, despite similar pre-licking behavior and velocity  
423 changes (Fig. 8e). This observation ruled out potential contributions of movement changes to  
424 the fast bi-phasic profile in the aDS GluCIN signal. Following light cue partial extinction (from  
425 100 to 20% reward probability), there was no significant change in this bi-phasic profile (Fig. 8e,  
426 early peak:  $p = 0.96$ , two-tailed Wilcoxon rank-sum test; late peak:  $p = 0.34$ , two-tailed Wilcoxon  
427 rank-sum test) indicating that an increase in glutamate release onto the CINs was not  
428 responsible for the increase of aDS ACh release during partial extinction (Fig. 4). Despite this,  
429 the latency of the slower component of the GluCIN release aligned, on average, with the latency  
430 of the ACh late peak increase following extinction (median latency to Glu late peak = 0.64s, vs  
431 median latency to ACh late peak = 0.8s), indicating that the influence of this input may be  
432 'unmasked' to drive the ACh increase, perhaps by disinhibition through emergent DA dips (see  
433 Discussion).

434 **DISCUSSION**

435

436 We applied high density arrays of small diameter optical fibers to map the spatiotemporal  
437 dynamics of rapid changes in ACh release across the striatum during Pavlovian learning and  
438 partial extinction in behaving mice. Our measurements reveal a topographic organization in the  
439 evolution of ACh release across timescales ranging from 100s of milliseconds to days as  
440 Pavlovian associations are learned then broken. Most notably, we identified a specific region of  
441 the anterior dorsal striatum (aDS) with changes in ACh release over learning and extinction  
442 consistent with positive and negative RPE encoding for cues, rewards, and unrewarded  
443 consummatory actions. Changes were inverse in polarity to those widely reported for DA  
444 release and RPE encoding was preferentially expressed in long latency components (dips and  
445 late peaks) of the multi-phasic ACh release profiles. Some changes in early peaks consistent  
446 with positive RPE encoding similar to DA were observed<sup>22,39</sup>, but were more sparse,  
447 concentrated ventrally, and were primarily restricted to the cue period. We hypothesize that  
448 positive prediction error encoding in ACh dips defines plasticity windows over time and space  
449 which facilitate initial formation of Pavlovian associations. In contrast, negative prediction error  
450 encoding in ACh late peak elevations may play an active role in erasing learned associations  
451 which are no longer valid, resulting in down-regulated behavioral responding to previously  
452 valued cues. These signals may support a general role for striatal ACh in forms of behavioral  
453 flexibility identified in previous manipulation studies. While our study identified the aDS as a  
454 potential locus for downshifting learned cue-reward associations, different forms of flexibility  
455 may recruit changes in ACh in distinct striatal regions. Indeed, one study observed a correlation  
456 between ACh transients in the medial nucleus accumbens shell and cocaine place preference  
457 extinction<sup>29</sup> and another observed ACh increases during extinction of lever pressing behavior in  
458 the dorsal medial striatum<sup>33</sup>. Further, ACh manipulations can result in diverse, sometimes  
459 conflicting, effects across task paradigms and regions. Future surveys of large-scale ACh  
460 release will be necessary to clarify the specific ACh dynamics across striatal regions associated  
461 with different forms of behavioral flexibility.

462

463 Recordings of striatal CIN firing and ACh release have largely focused on changes occurring  
464 during initial associative learning. These studies have primarily reported peak and dip  
465 components to predictive cues and rewards only partially consistent with positive RPE  
466 encoding<sup>8,10–12,22,69</sup>. One study reported positive RPE encoding in elevations in CIN firing in the  
467 ventral, but not the dorsal striatum<sup>22</sup>. Our results show that changes to cues and rewards are  
468 topographically organized across distinct striatal axes for different signal components and that  
469 full positive RPE encoding is present predominantly for the dip component, is cue modality  
470 specific, and is concentrated in the aDS. These findings indicate that discrepancies across prior  
471 studies may be due, at least in part, to limited spatial sampling and that the three  
472 components of the multi-phasic ACh release profile (early peak, dip, late peak) are shaped by  
473 different underlying mechanisms or inputs, which vary across striatal regions. Several studies  
474 provide evidence that ACh dips are enhanced by (though not dependent on) inhibitory DA D2  
475 receptor signaling on CINs<sup>11,12,60</sup>, which may contribute to the positive RPE encoding inverse to  
476 DA in aDS dips. Recordings of midbrain DA neurons have reported full RPE encoding to  
477 conditioned cues and rewards across a large majority of the recorded population (particularly in

478 VTA<sup>35,37,55</sup>). However, some regional differences in striatal DA release (and positive RPE  
479 encoding) have been observed<sup>70-73</sup>, so it is possible that spatial variations in positive DA RPE  
480 signaling drive localized inverse RPE encoding in ACh release. Alternatively (or in addition), DA  
481 may exert a different relative influence in CIN activity across regions<sup>65</sup>.

482

483 What are the possible consequences of the spatially organized encoding of positive RPEs? ACh  
484 dips have been suggested to open a plasticity window to permit DA elevations to drive  
485 potentiation of excitatory synapses onto D1 receptor expressing projection neurons during  
486 associative learning<sup>3,20,21,74,75</sup>. Our results suggest that ACh dips may gate plasticity as a function  
487 of learning selectively within specific regions of the dorsal striatum. Changes in both peaks and  
488 dips were expressed across partly non-overlapping regions for tone and light cues (Figs. 2m,  
489 S2g), suggesting that plasticity gating may be sensory modality specific across regions, perhaps  
490 aligning with the topography of glutamatergic inputs to striatal CINs<sup>53,54,65</sup>. We propose that  
491 regional variations in ACh release shape where, when, and how synaptic plasticity is expressed  
492 in striatal sub-circuits and cell types.

493

494 As cues lost their reward predictive value during extinction, we observed an elevation of late  
495 peaks concentrated in the aDS, the same region in which positive RPE encoding was observed  
496 in ACh dips (Figs. 4 and S4). Late peak changes were consistent with negative RPE encoding,  
497 representing the difference between the previously learned value of the predictive cue and the  
498 lower predictive value after extinction. Changes over learning also occurred in late ACh peaks to  
499 spontaneous unrewarded ITI licks, consistent with negative RPEs and partially overlapping with  
500 the aDS negative RPE encoding at cue (Figs. 5 and S6). These results suggest that putative  
501 positive and negative RPEs are encoded preferentially in the aDS in the dip and late peak  
502 components respectively. How might the negative RPE encoding emerge in ACh late peaks?  
503 Our evidence suggests that DA dips in the aDS may be converted to ACh peak elevations via  
504 the cessation of a tonic D2 receptor mediated inhibition of CINs. Consistent with this idea,  
505 optogenetic inhibition of DA release is capable of rapidly (<200ms) elevating ACh release  
506 *in-vivo*, an effect which depends on D2 receptors<sup>60</sup>. The timing of this effect is consistent with  
507 the average latency of ACh peaks relative to the DA dips (Fig. 7, albeit across separate groups  
508 of mice). Our glutamate release measurements onto CINs indicated a long-latency component  
509 of excitatory glutamate release which does not increase during extinction, but which may be  
510 ‘unmasked’ by the emergent DA dips (Fig. 8). These results implicate a conjunction of  
511 appropriately timed glutamate input and disinhibitory DA dips to drive an emergence of ACh  
512 peaks encoding negative RPEs. This may occur either through post-synaptic plasticity of  
513 glutamatergic inputs onto CINs<sup>18</sup>, or through immediate effects of the DA dips on elevating CIN  
514 excitability.

515

516 Why would the emergence of negative RPE encoding in ACh late peaks be spatially restricted  
517 to the aDS? One possibility is that the DA dips, encoding negative RPEs, which disinhibit the  
518 CINs are also restricted to aDS. Variations in negative RPE encoding have been observed  
519 across DA neurons and in DA release, though region specific DA recordings have not localized  
520 dips exclusively to the aDS<sup>37,71,73</sup>. A second possibility is that the glutamatergic input unmasked  
521 by the DA dip is aDS specific. In support of this, significant DA dips were present for both the

522 tone and light cues in the aDS, but late peaks in glutamate release were selective for the light.  
523 Similarly, ACh late peaks for the tone were not as prominent in the aDS and were shifted  
524 posterior and lateral relative to the light (Fig. 4l and S4c), indicating that distinct patterns of  
525 glutamate release may shape the regional topography of negative RPE encoding for different  
526 cue modalities. Third, DA may exert a stronger inhibitory influence on ACh in the aDS than in  
527 other regions. In support of this, studies in brain slices have observed a gradient of D2 mediated  
528 inhibition of CINs in the striatum, with particularly strong suppression in the aDS, and an  
529 opposite effect in the ventral striatum<sup>65</sup>. Simultaneous, striatum wide measurements of DA and  
530 ACh during learning and extinction will be needed to further investigate these potential  
531 mechanisms.

532

533 Extinction has been proposed to occur both through an ‘unlearning’ process involving reversal  
534 of plasticity occurring during associative acquisition and through new inhibitory learning<sup>64,76</sup>. The  
535 ACh peaks in the aDS are well positioned to promote both of these processes in parallel  
536 through their differential effects of synaptic plasticity on striatal projection neuron cell types.  
537 Pavlovian associations are believed to involve the strengthening of synapses onto D1 receptor  
538 expressing direct pathway neurons (dSPNs) driven by phasic elevations in DA to unpredicted  
539 rewards and reward-associated cues (e.g. positive RPEs)<sup>21,55,74</sup>. Studies *in-vitro* have found that  
540 M4 muscarinic ACh receptors, expressed preferentially on direct pathway spiny projection  
541 neurons (dSPNs), promote long term synaptic depression<sup>77,78</sup>. Thus, the increases in cue  
542 evoked ACh during extinction could weaken the previously potentiated synapses onto dSPNs.  
543 Importantly, while DA dips encoding negative RPEs have been proposed to contribute to  
544 extinction, it is unclear how DA dips alone could weaken previously potentiated dSPN synapses  
545 because low affinity D1 receptors are not strongly activated by basal DA levels and therefore  
546 are not sensitive to DA dips<sup>79</sup>. In contrast, CINs express the high affinity inhibitory D2 DA  
547 receptor, so DA dips promote rapid ACh increases through disinhibition<sup>60,79</sup>. Therefore,  
548 disinhibition of CINs may be a mechanism to convert negative prediction errors encoded by DA  
549 dips into ACh signals that can drive de-potentiation of dSPN synapses and ‘unlearning’ of  
550 invalid cue-reward associations. In parallel, the aDS ACh elevations (and DA dips) during  
551 extinction may elevate excitability (relative to dSPNs) and strengthen synapses of D2  
552 expressing indirect pathway neurons (iSPNs) via M1 receptors<sup>3,4,80</sup>. This may result in new  
553 inhibitory learning, where iSPNs become more strongly activated to the extinction cue to actively  
554 suppress appetitive responding.

555

## 556 METHODS

### 557 Mice

558 For all experiments adult male and female (n = 33 mice in total, postnatal 3-5 months, 24-30g),  
559 wild type (WT) C57BL6 /J (n=15, 10 males and 5 females, Jackson Labs, strain #00664) and  
560 heterozygous ChAT-cre mice (n=18, 10 males and 8 females, ChAT-IRES-Cre, Jackson Labs,  
561 strain# 006410) were used in this study. Mice were initially housed in groups and then  
562 individually housed following surgery, under standard laboratory conditions (20-26°C, 30-70%  
563 humidity; reverse 12-h/12 h light/dark cycle; light on at 9 p.m.) with *ad libitum* access to food  
564 and water, except during water scheduling. The number of mice used for each experiment and

565 analysis is indicated in figure legends and main text. Experiments were conducted during the  
566 dark cycle. All animal care and experimental procedures were performed in accordance with  
567 protocols approved by the Boston University Institutional Animal Care and Use Committee  
568 (protocol no. 201800554) and in compliance with the Guide for Animal Care and Use of  
569 Laboratory Animals.

## 570 Multi-fiber array fabrication, calibration, and micro-CT scanning for fiber localization

571 *Array fabrication:* The multifiber arrays to enable large-scale measurements of ACh (and DA)  
572 release across the 3D volume of the striatum were fabricated in-house as previously described<sup>48</sup>  
573 (Fig. 1a). Briefly, 51-99 optical fibers (50µm, 46µm core and 4µm cladding with 0.66NA, Fiber  
574 Optics Tech) were threaded into holes (55-60µm diameter) within a custom 3D-printed grid  
575 (3mm W x 5mm L, Boston Micro Fabrication) and secured with UV glue (Norland Optical  
576 Adhesive 61). Fibers were fixed at preset lengths from the bottom of the grid (measured under a  
577 dissection scope) and neighboring fibers were separated by at least 250 µm axially and 220 µm  
578 radially to target different locations throughout the striatum volume, ensuring no (or very  
579 minimal) overlap between fluorescence collection volumes <sup>47,81,82</sup>. Fiber ends on the  
580 non-implanted side were bundled and glued inside a polyimide tube (0.92mm ID, MicroLumen)  
581 and then cut with a sharp razor blade to a ~ 1cm length. The bundled fibers were then polished  
582 first on fine 6µm grained polishing paper followed by 3µm (ThorLabs) to create a uniform and  
583 smooth imaging surface and enable efficient light transmission through each fiber (Fig. 1a).

584 *Fiber localization:* Following the array fabrication, a calibration procedure was performed to  
585 match individual fibers on the implanted side to their corresponding locations on the bundle  
586 surface as previously described<sup>48</sup>. At the end of neuronal recordings and behavioral  
587 experiments, mice were injected intraperitoneally with Euthasol (400 mg kg-1, Covertus  
588 Euthanasia III) then perfused intracardially with phosphate-buffered saline (PBS 1%, Fisher  
589 Scientific), followed by paraformaldehyde (4% in 1% PBS, Fisher Scientific). The mice were  
590 decapitated, and the ventral side of the skull removed to expose the ventral side of the brain.  
591 The intact implanted brains were post-fixed for 24h in 4% paraformaldehyde, rinsed with 1%  
592 PBS then transferred to a Lugol iodine solution (Carolina Scientific) diluted 1:3 in distilled water  
593 for 4 to 6 days. The diffusion of the Lugol solution to the brain enhances tissue contrast for  
594 computerized tomography scanning to enable fiber localization. The CT scanning and fiber  
595 localization were previously described<sup>48</sup>. Briefly, the 3-D CT scans of the intact implanted brains  
596 were registered to the The Allen Mouse Brain Common Coordinate Framework atlas via a  
597 semi-automated landmark-based approach<sup>49,83</sup>. The fibers, bright in the CT scan, were then  
598 automatically identified via intensity thresholding, and the recording locations (ventral-most  
599 point) were mapped to their corresponding locations on the implanted grid and subsequently to  
600 their locations on the imaging surface. In addition to bringing all mice into a common coordinate  
601 space, registration with the atlas also enabled the recording locations to be automatically  
602 assigned atlas anatomical labels, which were further manually verified (Fig. 1d).

603 **Stereotaxic viral injections and chronic optical fiber implants**

604 Mice were anesthetized with isoflurane (3-4%) and placed in a stereotaxic frame (Kopf  
605 instruments) on an electric heating pad (Physitemp instruments) and administered  
606 buprenorphine extended release for pre-operative analgesia (3.25 mg kg<sup>-1</sup> subcutaneous,  
607 Etiqa XR). Following induction, isoflurane was held at 1-2% (in 0.8-1 L min<sup>-1</sup> pure oxygen) and  
608 body temperature maintained at 37°C, throughout the surgical procedure. The scalp was  
609 shaved and cleaned with iodine solution prior to exposing the skull. For experiments to record  
610 extracellular acetylcholine release using multi-fiber arrays, a large craniotomy was performed  
611 over the right hemisphere with a surgical drill (Midwest Tradition 790044, Avtec Dental RMWT)  
612 to expose the brain surface from -2.3 to 2 mm in the anterior-posterior (AP) direction and from  
613 0.49 to 3.4 mm in the medial-lateral (ML) direction relative to bregma. AAV9-hSyn-ACh3.0 (WZ  
614 Biosciences<sup>45</sup>), 2.07x10<sup>13</sup> GC ml<sup>-1</sup> diluted 1:2 in PBS was pressure-injected into the striatum of  
615 WT mice (n = 8) through a pulled glass pipette (tip diameter 30-50 µm) at 20-40 separate  
616 striatum locations chosen to maximize expression around fiber tips (200 nl at each location at a  
617 rate of 100 nl/min). For control experiments, a 1:2 mixture of AAV9-hsyn-ACh3.0-mut<sup>45</sup> (WZ  
618 Biosciences), 2.54x10<sup>13</sup> GC ml<sup>-1</sup> was injected into the striatum of WT mice (n = 3) using the  
619 same strategy. For experiments to record extracellular DA release using multi-fiber photometry,  
620 AAV5-CAG-dlight1.3b<sup>67</sup> (Addgene, # 111067), 1.7x x10<sup>13</sup> GC ml<sup>-1</sup> diluted 1:3 in PBS was injected  
621 into the striatum of WT mice (n = 4) at 10-40 total locations (200-800nl at each location) using  
622 the same procedure. Following injections, the multi-fiber array was mounted onto the stereotaxic  
623 manipulator, the dura gently removed, and the array slowly lowered into position. The  
624 craniotomy was sealed with a thin layer of Kwik-Sil (WPI), and the array was secured to the  
625 skull surface using Metabond (Parkell). To allow head fixation, a metal head plate and ring  
626 (Atlas Tool and Die Works) were next secured to the skull with Metabond, and the implant  
627 surface was covered with a mixture of Metabond and carbon powder (Sigma Aldrich) to reduce  
628 optical artifacts. The fiber bundle was protected by a cylindrical plastic tube, extending ~ 1-2 mm  
629 above the fiber bundle, and secured around the bundle using a mixture of Metabond and carbon  
630 powder.

631 To drive the suppression of ACh release from cholinergic interneurons with tetanus toxin light  
632 chain (TelC, Fig. 6), small, circular craniotomies were drilled bilaterally above the injection sites  
633 (from bregma, in mm; AP: 1, ML: ± 1.4). Then either  
634 pAAV2/8-hSyn-FLEX-TeLC-P2A-EYFP-WPRE (Addgene, #135391,<sup>84</sup>, a gift from Bernando  
635 Sabatini's Lab, 5.14x10<sup>13</sup> GC ml<sup>-1</sup>) or  
636 ssAAV-5/2-hSyn1-dlox-TeTxLC\_2A\_NLS\_dTomato(rev)-dlox-WRPE-hGHP (Viral Vector Facility  
637 University of Zurich, 4.1x10<sup>12</sup> VG ml<sup>-1</sup>) diluted 1:1 in PBS was bilaterally injected in the anterior  
638 dorsal medial striatum of ChAT-cre mice (n = 6) at 4-12 sites per hemisphere (300nl/site at a  
639 rate of 100nl/min) at the following coordinates in mm; AP: 0.8, ML: ± 1.25, DV: -2.5 and -3;  
640 AP:1, ML: ±1.4, DV:-2.75 and -3. Control ChAT-cre mice (n = 8 Chat-cre mice) were injected  
641 with saline using the same strategy. To validate that TelC expression in the aDS leads to a  
642 reduction in ACh release, two ChAT-cre mice were bilaterally injected in the aDS using the same  
643 coordinates with ssAAV-5/2-hSyn1-dlox-TeTxLC\_2A\_NLS\_dTomato(rev)-dlox-WRPE-hGHP  
644 diluted 1:1 in PBS. Additionally, the left hemisphere of each mouse was co-injected with

645 AAV9-hSyn-ACh3.0 diluted 1:2 in PBS. Following the injections, the craniotomies were sealed  
646 with Kwik-Sil (WPI), and the skull was sealed with Metabond (Parkell) and a metal head plate.

647 To measure extracellular glutamate release into cholinergic interneurons (Fig. 8), craniotomies  
648 were drilled above the injection sites in the right hemisphere (from bregma, in mm; AP: 1, ML:  
649 1.4). A 1:1 mixture in PBS of the genetically encoded glutamate sensor  
650 AAV9.hSyn-FLEX.8F-iGluSnFR.A184S<sup>68</sup> (Adgene, #106174), 1.8x10<sup>13</sup> GC ml<sup>-1</sup> was injected in  
651 aDS of ChAT cre mice (n = 2) at 6 sites (300nl/site at a rate of 100nl/min) at the following  
652 coordinates in mm: AP: 0.8, ML: 1.5, DV: -2.75, -3.25 and -3.75; AP: 1.1, ML: 1.5, DV: -2.75,  
653 -3.25 and -3.75. Then, a 100 µm core diameter optical fiber (MFC\_100/125- 0.37NA) attached  
654 to a zirconia ferrule (Doric) was slowly lowered into the medial region of the aDS (AP:1, ML:1.4)  
655 to a final depth of 3 mm from bregma. The craniotomies were sealed with Kwik-Sil (WPI), the  
656 optical fiber and a head plate were secured to the skull with Metabond (Parkell). After the  
657 surgeries, mice were placed in a cage with a heating pad and received postoperative injections  
658 of meloxicam (5 mg kg<sup>-1</sup> subcutaneous, Covertus) and 1 mL of saline per day subcutaneously  
659 for 4 days after surgery. Mice were individually housed and allowed to recover in their cages for  
660 at least 2 weeks after surgery.

### 661 Pavlovian conditioning task and behavior setup

662 *Behavioral apparatus:* One week before starting the Pavlovian conditioning training and  
663 photometry recordings, mice were placed on a water schedule, receiving 1 ml of water per day,  
664 and were maintained at 80-85% of their initial body weight for the duration of the experiments.  
665 Three to four days prior to training, mice were habituated to head fixation on the spherical  
666 styrofoam treadmill<sup>52</sup> (Smoothfoam, 8in diameter, Fig. 1b). The behavioral setup has been  
667 described in detail previously<sup>48</sup>. Briefly, mice were free to locomote on the treadmill during all  
668 experiments and the ball rotation in pitch, yaw, and roll directions was measured using optical  
669 computer mice (Logitech G203) through an acquisition board (NIDAQ, PCIe 6343). Water  
670 rewards (5µL/reward) were dispensed through a water spout operated by an electronically  
671 controlled solenoid valve (#161T012, Neptune Research), mounted on a post a few mm away  
672 from the mice's mouths. Tongue spout contacts (a proxy for licking) were monitored by a  
673 capacitive touch circuit connected to the spout and confirmed with live video taken from a  
674 camera positioned to capture orofacial movements (Blackfly S USB3, BFS-U3-16S2M-CS,  
675 Teledyne Flir).

676 *Pavlovian conditioning and partial extinction:* Approximately three weeks post-implantation,  
677 mice began training on a dual cue delay Pavlovian conditioning task (Figs. 1b, 2a and 4a). In  
678 each session (one session/day), mice received 40 presentations of two different cues in a  
679 pseudorandom order: light and tone (20 presentations of each). Light cues were presented via a  
680 LED (Thor labs, M470L3, 470 nm) calibrated to deliver light at 7mW intensity and mounted on a  
681 post holder ~ 20 cm away from the mouse, positioned 45 degrees contralateral to the implanted  
682 side. Tone cues (12 kHz, 80dB) were presented via a USB speaker placed ~ 30 cm from the  
683 mouse. Each cue was presented for 6s and was paired with a water reward (5µL) delivered with  
684 100% probability after a fixed 3s delay from the cue onset. An ITI was randomly drawn from a  
685 uniform distribution of 4-40s. A total of eight random non-contingent rewards per session were

686 delivered during the ITI periods. The mice were trained for 7-12 consecutive days until they  
687 learned that both light and tone cues were associated with the delivery of a water reward (as  
688 measured by the lick index, see below). They were then trained for an additional 2-6 days.  
689 Following initial learning, mice were submitted to a partial extinction phase (Fig. 4a) in which the  
690 reward probability associated with one of the two cues was downshifted to 20% (10% for DA  
691 experiments, Fig. 7), and the other cue to 80% (90% for DA experiments). During the extinction  
692 phase, mice received 60 presentations of the two cues (30 presentations of each). Training  
693 continued until mice showed significantly diminished pre-licking for the 20% cue relative to the  
694 80% for 4-7 sessions, then cue probabilities were reversed. The order of light and tone cue  
695 probabilities was counterbalanced across mice. For TelC experiments, TelC and control mice  
696 were trained for a maximum of 8 sessions for each extinction phase.

## 697 Multi-fiber photometry recordings

698 Fluorescence measurements from the multi-fiber arrays were conducted using a custom built  
699 microscope (Fig.1b) mounted on a 4' W x 8' L x 12" thick vibration isolation table (Newport).  
700 Details of the microscope were described previously<sup>48</sup>. Excitation light for the fluorescent  
701 sensors (ACh3.0, ACh4.3 mut, iGluSnFR and dLight 1.3b) was provided by two high power  
702 LEDs (470nm and 405nm; Thor labs, No. SOLIS-470C, SOLIS-405C). Excitation light was  
703 bandpass filtered (Chroma No. ET405/10 and ET473/24) then coupled into a liquid light guide  
704 (Newport No. 77632) with lenses (f = 60mm and 30mm, Thor labs No. LA1401-A and LA1805)  
705 and a collimating beam probe (Newport No. 76600). The liquid light guide was connected to a  
706 filter cube on the microscope, directing excitation light into the back aperture of the microscope  
707 objective (10x, 0.3NA, Olympus Model UPLFLN10X2) via a dichroic beam splitter (Chroma  
708 Model 59009bs). The light power at the focal plane of the objective was adjusted to be within  
709 the range of 80-85mW, resulting in a power of 1.6~2mW/mm<sup>2</sup> at the fiber tips<sup>48</sup>. Emission light  
710 was bandpass filtered (Chroma, No 525/50m) and focused with a tube lens (Thor labs, No  
711 TTL165-A) onto the CMOS sensor of the camera (Hamamatsu, Orca Fusion BT Gen III),  
712 creating an image of the fiber bundle (Fig. 1a). To enable precise manual focusing, the  
713 microscope was connected to a micromanipulator (Newport Model 96067-XYZ-R) and mounted  
714 on a rotatable arm extending over the head-fixation setup to facilitate positioning of the objective  
715 above the imaging surface over the mouse head. Imaging data acquisition was performed using  
716 HClImage live (HClImage live, Hamamatsu). Single wavelength excitation was carried out with  
717 continuous imaging at 30Hz (33.33ms exposure time), via internal triggering. Dual wavelength  
718 excitation was performed in a quasi-simultaneous externally triggered imaging mode, where the  
719 two LEDs were alternated and synchronized with imaging acquisition via 5V digital TTL pulses.  
720 470 nm excitation was alternated with 405 nm excitation at either 36Hz (20ms exposure time) or  
721 22Hz (33.33ms exposure time) to achieve a frame rate of 18 Hz or 11 Hz for each excitation  
722 wavelength, respectively. Recordings acquired at different sampling rates were downsampled or  
723 upsampled using a 1-D interpolation with Matlab's *interp1* function using the spline method. A  
724 custom MATLAB software controlled the timing and duration of TTL pulses through a  
725 programmable digital acquisition card (NIDAQ, National Instruments PCIe 6343). Voltage pulses  
726 were transmitted to the NIDAQ from the camera following the exposure of each frame to confirm  
727 proper camera triggering and to synchronize imaging data with behavior data.

## 728 Statistics and reproducibility

729 Data were processed and analyzed using built-in and custom functions in Matlab (Matworks,  
730 version 2020b, 2022b and 2023a), Python, or GraphPad Prism10 (GraphPad Software). Some  
731 fibers were excluded from analysis based on localization outside the striatum or poor  
732 signal-to-noise ratio (see below for details). Exclusion was performed prior to any statistical  
733 analysis of task related signals. Tests for significance are indicated in the text and figure  
734 legends. Sample sizes were chosen to effectively measure experimental parameters while  
735 remaining in compliance with ethical standards to minimize animal usage. There was no  
736 randomization or blinding conducted.

## 737 Multi-fiber photometry signal preprocessing

738 The acquired time series videos of the fiber bundles were first motion-corrected using a  
739 whole-frame cross-correlation algorithm described previously<sup>52,85</sup> then visually inspected to  
740 confirm post-correction image stability. Circular regions of interest (ROI, ~ 25µm diameter) were  
741 manually selected for each fiber. The resulting set of ROIs comprised a mouse-specific ROI  
742 template, which was then fit and applied to each subsequent imaging video, enabling the  
743 identity of each ROI to remain consistent across multiple recording sessions. To determine the  
744 change in fluorescence ΔF/F, the mean fluorescence extracted from each ROI was normalized  
745 to a baseline, which was defined as the 8<sup>th</sup> percentile fluorescence over a 30-s sliding window<sup>48</sup>.  
746 To remove low frequency artifacts, the ΔF/F signals were high-pass filtered using a finite  
747 impulse response filter with a passband frequency set at 0.3Hz. This frequency was determined  
748 based on the observed differences in the dynamics of the ACh signal compared to the control  
749 signals (ACh-mut and the isosbestic 405 nm LED signal, Fig. 1h,i). Most analyses were  
750 conducted on non z-scored or peak normalized ΔF/F values in order to identify relative  
751 differences in signal magnitude across task and training phases. Changes in overall signal  
752 magnitude over training due to changes in sensor expression or fiber collection efficiency were  
753 accounted for by examining the stability of signals to unpredicted water reward for each fiber  
754 (Fig.1f).

## 755 Quantification and statistical analysis

### 756 Definition of learning phases

757 Lick indices were computed for each trial and session to assess learning of the Pavlovian  
758 associations. The lick index was defined by the following formula:

$$759 \text{Lick index} = \max(0, \log(\text{anticipatory lick}) \times ((\text{anticipatory lick} - \text{lick ITI}) \div (\text{anticipatory lick} + \text{lick ITI}))$$

760 where anticipatory lick is the sum of lick spout counts across a 1s window before reward  
761 delivery, and lick ITI is the sum of lick count across a 1s window before cue onset. To determine  
762 the learning phases (pre- and post-learning), the mean lick indices across all trials of each  
763 acquisition session were compared to those of the first acquisition session using a two tailed  
764 Wilcoxon rank-sum test. A p-value <0.05 was considered statistically significant. For initial  
765 Pavlovian conditioning, sessions where the mean lick indices for both cues were not

766 significantly different from the first session were considered pre-learning sessions ( $2.9 \pm 0.62$ ),  
767 while sessions where mean lick indices of both cues were significantly higher than the first  
768 session were considered post-learning sessions ( $3.27 \pm 0.5$ ) (Fig. 2c,d). Sessions where only one  
769 cue had a significantly elevated lick index were omitted from analysis ( $2.90 \pm 0.75$  sessions). For  
770 the extinction phase (Fig. 4), sessions were included in analyses if the lick indices of the 20%  
771 probability cue were significantly (Wilcoxon rank-sum test,  $p < 0.05$ ) lower than lick indices for the  
772 same cue on the last day of the post-learning phase or the preceding 80% probability session  
773 (for reversal sessions). The fraction of time spent licking during the cue period (Fig. 2d) was  
774 calculated across the entire 3-s window after cue onset, and the fraction of time spent licking  
775 during the ITI period was calculated across the entire ITI period, excluding 0.5s before cue  
776 onset, 3s after cue offset, and 6s after any unpredicted reward delivery. Spontaneous lick bout  
777 onsets in the ITI (Figs. 5 and 7) were defined as the first lick of lick bouts that are not preceded  
778 by any licking for at least two seconds.

779 For lick analyses in TelC experiments (Fig. 6), all the extinction sessions (20% reward  
780 probability for both light and tone cues combined) were compared with sessions of pre- and  
781 post-learning (Two-way ANOVA followed by Tukey *post hoc* analysis to account for multiple  
782 comparisons within and between TelC and control mice,  $p < 0.05$  was considered statistically  
783 significant).

#### 784 Relationships of signal with velocity and acceleration

785 Analog signals from the optical mice were converted to m/s<sup>48</sup>, and the pitch and roll were  
786 combined to compute a total velocity calculated as:  $\sqrt{pitch^2 + roll^2}$ . The velocity traces were  
787 then smoothed using a Savitzky-Golay moving average filter with a moving window of 250 ms  
788 and a 2nd degree polynomial. Acceleration traces were derived from the smoothed velocity  
789 (acceleration =  $\Delta$ velocity/ $\Delta$ time), and further filtered using the same Savitzky-Golay moving  
790 average filter parameters.

791 To partially account for generalized relationships between treadmill locomotion and  $\Delta F/F$ , the  
792 filtered fluorescence signals during ITI periods were fit to a generalized linear model (GLM)  
793 using smoothed linear velocity and acceleration as continuous predictors, each with positive and  
794 negative phases. The optimal positive and negative phase differences between ITI  $\Delta F/F$  and  
795 velocity/acceleration were first identified through cross-correlations with a maximum lag of  
796  $\pm 0.5$ s. Next, correlation coefficients were calculated by fitting the phase-shifted  
797 velocities/accelerations to the ITI  $\Delta F/F$  via least squares linear regression using Matlab's *fitglm*  
798 (equations below):

#### 799 Equation 1:

$$800 signal_{\Delta F/F_{filtered}} = \beta_0 + \beta_{v1} \times v_1 + \beta_{v2} \times v_2 + \beta_{a1} \times a_1 + \beta_{a2} \times a_2 + \epsilon$$

801 GLM training, where  $signal_{\Delta F/F_{filtered}}$  is the filtered  $\Delta F/F$  of the ITI periods,  $v_1/v_2$  and  $a_1/a_2$  are  
802 positive/negative phase-shifted velocities and accelerations respectively,  $\beta_{v1}/\beta_{v2}$  are the

803 correlation coefficients of positive/negative phase-shifted velocities,  $\beta_{a1}/\beta_{a2}$  are the correlation  
804 coefficients of positive/negative phase-shifted accelerations,  $\varepsilon$  is the error term.

805

806 Equation 2:

$$807 \text{signal}_{velocity} = \widehat{\beta}_0 + \widehat{\beta}_{v1} \times v_1 + \widehat{\beta}_{v2} \times v_2 + \widehat{\beta}_{a1} \times a_1 + \widehat{\beta}_{a2} \times a_2$$

808  $\text{signal}_{velocity}$  is the estimated contribution of velocity and acceleration to the  $\Delta F/F$  signal,  $v_1/v_2$

809 and  $a_1/a_2$  are the same phase-shifted velocities and accelerations as described above,  $\widehat{\beta}$  are

810 the estimated correlation coefficients generated by the GLM. Non-significant  $\widehat{\beta}$ , i.e. with a  $p >$   
811 0.05, were set to 0 and were not included in the estimated corresponding velocity or  
812 acceleration contribution.

813 The final  $\Delta F/F$  was computed by subtracting the velocity contribution ( $\text{signal}_{velocity}$ ) from the  
814 filtered signal ( $\Delta F/F_{filtered}$ ).

815 To further address the possibility that variations in treadmill velocity at cues or rewards  
816 contribute to  $\Delta F/F$  signal changes, signals and velocity/acceleration were compared for different  
817 learning phases for each fiber (Fig. S7). To identify decelerations occurring to cues, rewards, or  
818 lick bout onsets, deceleration periods were first defined in peri-event windows as consecutive  
819 bins with negative acceleration values. Peak deceleration (Fig. S7a-c; Fig. 7g) was defined as  
820 the minimum value of acceleration during any single continuous deceleration period, in which  
821 the total velocity change exceeded a threshold of 8 cm/s. Total velocity changes following cue or  
822 reward (Fig. S7a-c; Fig. 7g) were defined as the maximum change of velocity following the  
823 event onset, relative to pre-event. Total velocity change and peak decelerations were calculated  
824 from cue onset to 0.6s after, and from -0.2 to 0.6s relative to reward consumption (first spout lick  
825 after delivery). For non-contingent spontaneous licks during ITI periods, total velocity change  
826 and peak decelerations were calculated from -0.5 to 0.1s after lick onset. Large and small  
827 deceleration trials were defined as trials where the peak deceleration was greater than the 80th  
828 percentile or less than the 20th percentile of the total peak deceleration distribution across all  
829 mice, respectively (Fig. S7h-k). For analyzing the contribution of signals related to spontaneous,  
830 non-task related decelerations (Fig. S7d-g), decelerations were first identified within ITI periods,  
831 excluding all times around spontaneous licks and unpredicted reward deliveries. Periods of  
832 continuous negative acceleration were identified and the troughs were included in triggered  
833 average analysis if they were lower than  $-1.5 \text{m/s}^2$ , and there was a total velocity change over  
834 the deceleration period of at least 20 cm/s. These thresholds were chosen conservatively to  
835 ensure that ITI decelerations matched (or exceeded) the average deceleration magnitudes  
836 observed in the task (Fig S7a-c). ACh  $\Delta F/F$  signals were aligned to the deceleration trough, and  
837 maximum and minimum values of the mean trough triggered average were calculated from a  $\pm$   
838 1s window around the peak. Statistical significance of minima and maxima was determined by  
839 comparison to the 95% confidence intervals of a bootstrap-generated random sampling  
840 distribution (2000 iterations). To evaluate the potential influence of decelerations occurring at the  
841 cue or reward on learning-related changes, correlations (Pearson's) were calculated between  
842 significant ITI deceleration minima or maxima and the magnitude of the mean  $\Delta F/F$  for different

843 signal components during the cue and reward periods, limited to fibers with both significant  
844 deceleration and task-related changes for each component (Fig. S7I,m; Table 1).

845 *Identification of multi-phasic ACh release components and changes with learning*

846 Analysis was conducted on 295 fibers across the 3D volume of the striatum, out of a total of 505  
847 implanted fibers, collected from 8 mice (with 37,36,47,44,45,31,29 and 26 fibers in each  
848 respective mouse). Additionally, analysis was performed on 110 fibers across the striatum, out  
849 of 185 implanted fibers, obtained from 3 mice expressing the non-functional ACh sensor (with  
850 43, 44, and 23 fibers in each mouse respectively).

851 Significant positive or negative changes in ACh release were first identified around each event  
852 (cue onset, reward delivery, lick bout onset) within peri-event windows (0-1.5s for cue, -0.5 -1.5s  
853 for reward, and -0.5-1s for lick onsets). Significance was determined as a signal change  
854 exceeding 3 standard deviations above (peaks) or below (dips) the mean of the ITI signal in a  
855 1s window prior to the event window. Local maximum (peaks) and minimum (dips)  $\Delta F/F$  points  
856 were defined for each fiber and the timing of these significant points was plotted in histograms  
857 (Fig. S1a,b; Fig. S6a). Based on the multi-phasic timing distributions from the histograms,  
858 windows were defined for each event to define significant early peak, dip, and late peak  
859 components (Fig. S1a,b). The windows were as follows for early peak, dip, and late peak  
860 respectively: cue onset: 55ms to 444ms, 333ms to 833ms, 500ms to 1222ms; reward  
861 consumption: -200ms to 400ms, 200ms to 800ms, 500ms to 1200ms; ITI lick bout onset:  
862 -270ms to 230ms, 0ms to 600ms, 200ms to 800ms. For maps showing the presence of each set  
863 of components for each fiber, additional criteria were included that a signal must exceed 3  
864 standard deviations relative to the ITI for at least 2 consecutive 0.056ms time bins and reach a  
865 peak or trough of at least 0.005  $\Delta F/F$  or -0.005  $\Delta F/F$  respectively to isolate only the largest  
866 signal changes.

867 Differences in the mean event-evoked  $\Delta F/F$  between different phases of learning for the three  
868 ACh signal components were calculated for each fiber by comparing the peak and dip values  
869 across single trials within each phase using an unpaired Wilcoxon rank sum test. *P*-values  
870 <0.01 were considered statistically significant, except for the difference in mean ITI lick  
871 bout-evoked  $\Delta F/F$ , where *p*-values <0.025 were considered significant. Signal changes  
872 consistent with RPE signaling were determined for each signal component and fiber across  
873 pairs of events (Figs. 3d,5g; Figs. S3e,S6f). Putative positive RPE encoding was defined as  
874 elevations in  $\Delta F/F$  (larger peaks or deeper dips) with learning to the cue and decreases to the  
875 reward. Negative RPE encoding was defined as elevations to the cue with extinction and  
876 decreases to unrewarded ITI licks with learning.

877 To visualize the spatial distribution of learning related ACh changes across the striatum, ACh  
878  $\Delta F/F$  values for each fiber, representing the mean difference between phases of learning, were  
879 aggregated into smoothed heat maps (e.g. Fig. 2j,k). First, mean  $\Delta F/F$  difference values were  
880 binned into cubic arrays with an edge length of 0.05 mm based on the spatial location of each  
881 fiber. Fibers with no significant difference or no significant component were included as 0s.  
882 Values from fibers within each spatial bin were averaged then convolved with an exponentially

883 decaying, distance-dependent spherical mask, with the decay rate calculated as: decay rate =  
884  $e^{-6 \cdot \text{euclidean distance}}$ . The result was a 3-dimensional array of weighted sums for each spatial location,  
885 which was then reduced to 2-dimensional sagittal or axial plane plots by averaging values along  
886 the collapsed dimension. Regions with the largest signal changes (white contours) were  
887 identified as containing points with amplitudes above the 90<sup>th</sup> percentile of the corresponding  
888 reduced 2D array for increases and decreases separately.

#### 889 Spatial gradient analysis

890 To quantify the variation in signal changes with learning/extinction along each anatomical axis  
891 (Figs 2l,m and 4k,l), a linear mixed-effects model was constructed which described the signal  
892 change amplitudes as a function of spatial coordinates along the anterior-posterior (AP),  
893 medial-lateral (ML), and dorsal-ventral (DV) axes. Spatial coordinates were included as fixed  
894 effects, while mouse identity was treated as a random effect to account for individual variability  
895 between subjects. The model is described by the following equation:

$$896 y_{ijk} = \beta_0 + \beta_{AP} \times AP_i + \beta_{ML} \times ML_j + \beta_{DV} \times DV_k + u_{AP, mouse} + u_{ML, mouse} + u_{DV, mouse} + u_{mouse} + \epsilon_{ijk}$$

897 where  $y_{ijk}$  is the amplitude of the ΔF/F differences,  $\beta_0$  is the intercept,  $\beta_{AP}$ ,  $\beta_{ML}$ ,  $\beta_{DV}$  are the fixed  
898 effect coefficients for the AP, ML and DV coordinates respectively,  $u_{mouse}$ ,  $u_{AP, mouse}$ ,  $u_{ML, mouse}$ ,  
899  $u_{DV, mouse}$  are random effect intercept and coefficients for AP, ML and DV coordinates, for  
900 individual mice, and  $\epsilon_{ijk}$  is the residual error term. [ $\beta_{AP}$ ,  $\beta_{ML}$ ,  $\beta_{DV}$ ] defined the striatal axes of  
901 maximal variation. The sign of the spatial coefficients was set so that the direction of the  
902 predominant signal change for each component was positive (e.g. larger dips or larger peaks  
903 were both positive), so that the sign of the coefficient indicated the direction of the spatial  
904 gradient (rather than the sign of the change) for comparison across components that changed in  
905 opposite directions.

906 To compare gradient coefficients between ACh components or experimental contexts (light vs  
907 tone for example), a similar model was fit, which included an interaction term between the  
908 coefficients of interest. This model is represented by the following equation:

$$909 y_{ijkl} = \beta_0 + \beta_{AP} \times AP_i + \beta_{ML} \times ML_j + \beta_{DV} \times DV_k + \beta_g \times Group_l + \beta_{g, AP} \times Group_l \times AP_i + \\ 910 \beta_{g, ML} \times Group_l \times ML_j + \beta_{g, DV} \times Group_l \times DV_k + u_{AP, mouse} + u_{ML, mouse} + u_{DV, mouse} + u_{g, mouse} + u_{mouse} \\ 911 + \epsilon_{ijk}$$

912 In this model,  $y_{ijkl}$  is the amplitude of the ΔF/F differences, constructed by stacking the two  
913 amplitude datasets for comparison, with their corresponding AP, ML and DV coordinates.  $Group_l$   
914 is a categorical variable distinguishing the two coefficients (e.g. light or tone) being compared,  
915  $\beta_g$  represents the coefficients of the categorical variable,  $\beta_{g, AP}$ ,  $\beta_{g, ML}$ ,  $\beta_{g, DV}$  are the interaction  
916 coefficients between groups and AP, ML, and DV,  $u_{g, mouse}$  are random effect coefficients for  
917 groups. All other parameters not explicitly mentioned are the same as the non-interaction

918 model. Wald t-test was used for statistical significance, a  $p$ -value  $< 0.05$  of the interaction  
919 coefficients i.e.  $\beta_{g, AP}$ ,  $\beta_{g, ML}$  and  $\beta_{g, DV}$  is considered significant

920

### 921 *Relative signal timing calculations*

922 Latencies for ACh peaks or dips were determined as the time between the event onset and the  
923 time of the minimum or maximum signal amplitude in the triggered average for a given signal  
924 component (early/late peak, dip). To determine the relative timing between the increase in light  
925 and tone cue-evoked late peak  $\Delta F/F$  and the subsequent decrease in lick index following partial  
926 extinction (Fig. 4m-p and 7f), lick indices and late peak amplitude values for each fiber across  
927 trials were smoothed using Matlab's *smooth* function with the lowess method (local regression  
928 using weighted linear least squares and a first-degree polynomial model with a moving window  
929 of 30 trials). Next, Matlab's *cusum* function with a climit of 5 and mshift of 2 was used to identify  
930 the change point for lick indices and late peak  $\Delta F/F$  (a decrease in lick index and an increase in  
931 late peak). The *cusum* algorithm detects small incremental changes by maintaining cumulative  
932 sums of detectable positive or negative shifts from the mean of each data point in the sequence.  
933 The threshold for detectable shifts was determined by the product of mshift and the standard  
934 deviation of the sequence. Significance was established when the upper or lower cumulative  
935 sum of shifts exceeded a threshold based on the product of climit and the standard deviation of  
936 the sequence.

### 937 *Dopamine changes during learning and extinction*

938 Measurements of DA release (dLight1.3b  $\Delta F/F$ , Fig. 7) were performed and analyzed similarly to  
939 ACh data. For the analysis of DA signals, 29 fibers in the aDS were selected from a total of 150  
940 fibers within the striatum (out of 324 implanted fibers), based on the criteria AP>0mm, ML<2mm  
941 and DV>-4mm. These fibers were selected from 4 mice (10, 6, 6, 7 fibers in each mouse  
942 respectively) implanted with arrays. Other fibers were ignored for this study and anatomical  
943 selection criteria was defined only with respect to the regional ACh signaling patterns in the  
944 ACh3.0 group and blindly with respect to the DA signals. Dips in DA release were identified as  
945 the minimum values of the trial averaged dLight1.3b  $\Delta F/F$  within a time window 0.2 to 1.5s after  
946 cue onset or spontaneous unrewarded lick bouts. The latencies of DA dips to spontaneous  
947 unrewarded licks bouts in the ITI were calculated as the time elapsed between the ITI lick bout  
948 onsets and the minimum  $\Delta F/F$ , averaged across trials (Fig. 7k). The latencies between the  
949 increase in cue-evoked DA dips and decrease in lick indices during extinction was determined  
950 using the same analysis as for ACh (see above, Fig. 7f).

### 951 *Histology*

952 At the end of glutamate recordings and TeIC behavior experiments, mice were injected  
953 intraperitoneally with Euthasol ( $400 \text{ mg kg}^{-1}$ , Covertus Euthanasia III) then perfused intracardially  
954 with phosphate-buffered saline (PBS 1%, Fisher Scientific), followed by paraformaldehyde (PFA  
955 4% in 1% PBS, Fisher Scientific). The brains were post-fixed overnight in 4% PFA dissolved in  
956 PBS and then transferred to a solution of 40% sucrose in 1% PBS (until the brains sank). The  
957 brains were sliced ( $50\mu\text{m}$  thickness) with a cryostat (Leica CM3050 S). The coronal sections

958 were then mounted on super frost slides and cover slipped with Vectashield antifade mounting  
959 medium (Vector Laboratories, H-1900). TeIC (Fig. 6b) and iGluSnFr (Fig. 8b) fluorescence were  
960 not immuno-enhanced; confocal images were acquired on a Zeiss LSM 800 laser scanning  
961 confocal microscope.

## 962 ACKNOWLEDGEMENTS

963

964 This work was supported by Aligning Science Across Parkinson's (ASAP, ASAP-020370)  
965 through the Michael J. Fox Foundation for Parkinson's Research (MJFF), Stanley Fahn Young  
966 Investigator Award (PF-SF-JFA-836662), Klingenstein-Simons's Foundation fellowship,  
967 Whitehall Foundation Fellowship, National Institute of Mental Health (R01 MH125835) to  
968 M.W.H; NIMH F32MH120894 to M.-A.T.V. We thank the Boston University Centers for  
969 Neurophotonics and Systems Neuroscience for financial and technical support, Micro CT core,  
970 especially Sydney Holder, for providing equipment and technical expertise for micro-CT  
971 scanning and Boston University Animal Science Center for providing central laboratory and  
972 animal care and support resources. We thank Glenda Smerin and Anosha Khawaja-Lopez for  
973 their technical help. We thank Jason Climer and Daniel Dombeck for the motion correction  
974 algorithm code, Lynne Chantranupong and Bernardo Sabatini for providing the TeIC virus, and  
975 Nicolas Tritsch and Eleanor Brown for the helpful discussions of this work.

976

## 977 AUTHOR CONTRIBUTIONS

978 Conceptualization - MWH, SB; Methodology - SB, LZ, MTV, KT, BMG, CAN, MWH; Software -  
979 LZ, MTV; Formal Analysis - LZ, SB, MWH; Investigation - SB, MTV, KT, BMG, CAN; Data  
980 Curation - SB, LZ; Writing Original Draft - SB, LZ, MTV, MWH; Visualization - SB, LZ, MWH;  
981 Supervision - MWH; Funding Acquisition - MTV, MWH.

## 982 DECLARATION OF INTERESTS

983 The authors declare no competing interests.

## 984 REFERENCES

- 985 1. Bolam, J.P., Wainer, B.H., and Smith, A.D. (1984). Characterization of cholinergic neurons in  
986 the rat neostriatum. A combination of choline acetyltransferase immunocytochemistry,  
987 Golgi-impregnation and electron microscopy. *Neuroscience* 12, 711–718.  
988 [https://doi.org/10.1016/0306-4522\(84\)90165-9](https://doi.org/10.1016/0306-4522(84)90165-9).
- 989 2. Butcher, S.G., and Butcher, L.L. (1974). Origin and modulation of acetylcholine activity in the  
990 neostriatum. *Brain Research* 71, 167–171. [https://doi.org/10.1016/0006-8993\(74\)90202-9](https://doi.org/10.1016/0006-8993(74)90202-9).
- 991 3. Goldberg, J.A., Ding, J.B., and Surmeier, D.J. (2012). Muscarinic modulation of striatal  
992 function and circuitry. *Handb Exp Pharmacol*, 223–241.  
993 [https://doi.org/10.1007/978-3-642-23274-9\\_10](https://doi.org/10.1007/978-3-642-23274-9_10).
- 994 4. Lim, S.A.O., Kang, U.J., and McGehee, D.S. (2014). Striatal cholinergic interneuron  
995 regulation and circuit effects. *Front. Synaptic Neurosci.* 6.  
996 <https://doi.org/10.3389/fnsyn.2014.00022>.
- 997 5. Mallet, N., Leblois, A., Maurice, N., and Beurrier, C. (2019). Striatal Cholinergic Interneurons:  
998 How to Elucidate Their Function in Health and Disease. *Front Pharmacol* 10, 1488.  
999 <https://doi.org/10.3389/fphar.2019.01488>.
- 1000 6. Apicella, P. (2017). The role of the intrinsic cholinergic system of the striatum: What have we  
1001 learned from TAN recordings in behaving animals? *Neuroscience* 360, 81–94.  
1002 <https://doi.org/10.1016/j.neuroscience.2017.07.060>.
- 1003 7. Kimura, M., Rajkowski, J., and Evarts, E. (1984). Tonically discharging putamen neurons  
1004 exhibit set-dependent responses. *Proc. Natl. Acad. Sci. U.S.A.* 81, 4998–5001.  
1005 <https://doi.org/10.1073/pnas.81.15.4998>.
- 1006 8. Aosaki, T., Tsubokawa, H., Ishida, A., Watanabe, K., Graybiel, A.M., and Kimura, M. (1994).  
1007 Responses of tonically active neurons in the primate's striatum undergo systematic changes  
1008 during behavioral sensorimotor conditioning. *J. Neurosci.* 14, 3969–3984.  
1009 <https://doi.org/10.1523/JNEUROSCI.14-06-03969.1994>.
- 1010 9. Apicella, P., Legallet, E., and Trouche, E. (1997). Responses of tonically discharging neurons  
1011 in the monkey striatum to primary rewards delivered during different behavioral states: *Exp  
1012 Brain Res* 116, 456–466. <https://doi.org/10.1007/PL00005773>.
- 1013 10. Apicella, P., Deffains, M., Ravel, S., and Legallet, E. (2009). Tonically active neurons in  
1014 the striatum differentiate between delivery and omission of expected reward in a probabilistic  
1015 task context. *European Journal of Neuroscience* 30, 515–526.  
1016 <https://doi.org/10.1111/j.1460-9568.2009.06872.x>.
- 1017 11. Martyniuk, K.M., Torres-Herraez, A., Lowes, D.C., Rubinstein, M., Labouesse, M.A., and  
1018 Kellendonk, C. (2022). Dopamine D2Rs coordinate cue-evoked changes in striatal  
1019 acetylcholine levels. *eLife* 11, e76111. <https://doi.org/10.7554/eLife.76111>.
- 1020 12. Gallo, E.F., Greenwald, J., Yeisley, J., Teboul, E., Martyniuk, K.M., Villarin, J.M., Li, Y.,  
1021 Javitch, J.A., Balsam, P.D., and Kellendonk, C. (2022). Dopamine D2 receptors modulate the  
1022 cholinergic pause and inhibitory learning. *Mol Psychiatry* 27, 1502–1514.  
1023 <https://doi.org/10.1038/s41380-021-01364-y>.
- 1024 13. Aosaki, T., Graybiel, A.M., and Kimura, M. (1994). Effect of the Nigrostriatal Dopamine  
1025 System on Acquired Neural Responses in the Striatum of Behaving Monkeys. *Science* 265,  
1026 412–415. <https://doi.org/10.1126/science.8023166>.
- 1027 14. Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but  
1028 Distinct Messages of Midbrain Dopamine and Striatal Tonically Active Neurons. *Neuron* 43,  
1029 133–143. <https://doi.org/10.1016/j.neuron.2004.06.012>.
- 1030 15. Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and Bergman, H. (2008). Midbrain  
1031 Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between  
1032 Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning  
1033 Trials. *J. Neurosci.* 28, 11673–11684. <https://doi.org/10.1523/JNEUROSCI.3839-08.2008>.

- 1034 16. Ding, J.B., Guzman, J.N., Peterson, J.D., Goldberg, J.A., and Surmeier, D.J. (2010).  
1035 Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron* 67, 294–307.  
1036 <https://doi.org/10.1016/j.neuron.2010.06.017>.
- 1037 17. Kharkwal, G., Radl, D., Lewis, R., and Borrelli, E. (2016). Dopamine D2 receptors in  
1038 striatal output neurons enable the psychomotor effects of cocaine. *Proceedings of the  
1039 National Academy of Sciences* 113, 11609–11614. <https://doi.org/10.1073/pnas.1608362113>.
- 1040 18. Zhang, Y.-F., Reynolds, J.N.J., and Cragg, S.J. (2018). Pauses in Cholinergic  
1041 Interneuron Activity Are Driven by Excitatory Input and Delayed Rectification, with Dopamine  
1042 Modulation. *Neuron* 98, 918–925.e3. <https://doi.org/10.1016/j.neuron.2018.04.027>.
- 1043 19. Cragg, S.J. (2006). Meaningful silences: how dopamine listens to the ACh pause.  
1044 *Trends in Neurosciences* 29, 125–131. <https://doi.org/10.1016/j.tins.2006.01.003>.
- 1045 20. Reynolds, J.N.J., Avvisati, R., Dodson, P.D., Fisher, S.D., Oswald, M.J., Wickens, J.R.,  
1046 and Zhang, Y.-F. (2022). Coincidence of cholinergic pauses, dopaminergic activation and  
1047 depolarisation of spiny projection neurons drives synaptic plasticity in the striatum. *Nat  
1048 Commun* 13, 1296. <https://doi.org/10.1038/s41467-022-28950-0>.
- 1049 21. Calabresi, P., Picconi, B., Tozzi, A., and Di Filippo, M. (2007). Dopamine-mediated  
1050 regulation of corticostriatal synaptic plasticity. *Trends Neurosci* 30, 211–219.  
1051 <https://doi.org/10.1016/j.tins.2007.03.001>.
- 1052 22. Duhne, M., Mohebi, A., Kim, K., Pelattini, L., and Berke, J.D. (2024). A mismatch  
1053 between striatal cholinergic pauses and dopaminergic reward prediction errors. *Proc. Natl.  
1054 Acad. Sci. U.S.A.* 121, e2410828121. <https://doi.org/10.1073/pnas.2410828121>.
- 1055 23. Aoki, S., Liu, A.W., Zucca, A., Zucca, S., and Wickens, J.R. (2015). Role of Striatal  
1056 Cholinergic Interneurons in Set-Shifting in the Rat. *J Neurosci* 35, 9424–9431.  
1057 <https://doi.org/10.1523/JNEUROSCI.0490-15.2015>.
- 1058 24. Aoki, S., Liu, A.W., Akamine, Y., Zucca, A., Zucca, S., and Wickens, J.R. (2018).  
1059 Cholinergic interneurons in the rat striatum modulate substitution of habits. *Eur J Neurosci*  
1060 47, 1194–1205. <https://doi.org/10.1111/ejn.13820>.
- 1061 25. Bradfield, L.A., Bertran-Gonzalez, J., Chieng, B., and Balleine, B.W. (2013). The  
1062 thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with  
1063 existing learning in the striatum. *Neuron* 79, 153–166.  
1064 <https://doi.org/10.1016/j.neuron.2013.04.039>.
- 1065 26. Ragozzino, M.E. (2003). Acetylcholine actions in the dorsomedial striatum support the  
1066 flexible shifting of response patterns. *Neurobiol Learn Mem* 80, 257–267.  
1067 [https://doi.org/10.1016/s1074-7427\(03\)00077-7](https://doi.org/10.1016/s1074-7427(03)00077-7).
- 1068 27. Ragozzino, M.E., Mohler, E.G., Prior, M., Palencia, C.A., and Rozman, S. (2009).  
1069 Acetylcholine activity in selective striatal regions supports behavioral flexibility. *Neurobiol  
1070 Learn Mem* 91, 13–22. <https://doi.org/10.1016/j.nlm.2008.09.008>.
- 1071 28. Ragozzino, M.E., Jih, J., and Tzavos, A. (2002). Involvement of the dorsomedial striatum  
1072 in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res* 953, 205–214.  
1073 [https://doi.org/10.1016/s0006-8993\(02\)03287-0](https://doi.org/10.1016/s0006-8993(02)03287-0).
- 1074 29. Fleming, W., Lee, J., Briones, B.A., Bolkan, S.S., and Witten, I.B. (2022). Cholinergic  
1075 interneurons mediate cocaine extinction in male mice through plasticity across medium spiny  
1076 neuron subtypes. *Cell Reports* 39, 110874. <https://doi.org/10.1016/j.celrep.2022.110874>.
- 1077 30. Collins, A.L., Aitken, T.J., Huang, I.-W., Shieh, C., Greenfield, V.Y., Monbouquette, H.G.,  
1078 Ostlund, S.B., and Wassum, K.M. (2019). Nucleus Accumbens Cholinergic Interneurons  
1079 Oppose Cue-Motivated Behavior. *Biol Psychiatry* 86, 388–396.  
1080 <https://doi.org/10.1016/j.biopsych.2019.02.014>.
- 1081 31. Gangal, H., Xie, X., Huang, Z., Cheng, Y., Wang, X., Lu, J., Zhuang, X., Essoh, A.,  
1082 Huang, Y., Chen, R., et al. (2023). Drug reinforcement impairs cognitive flexibility by inhibiting  
1083 striatal cholinergic neurons. *Nat Commun* 14, 3886.  
1084 <https://doi.org/10.1038/s41467-023-39623-x>.

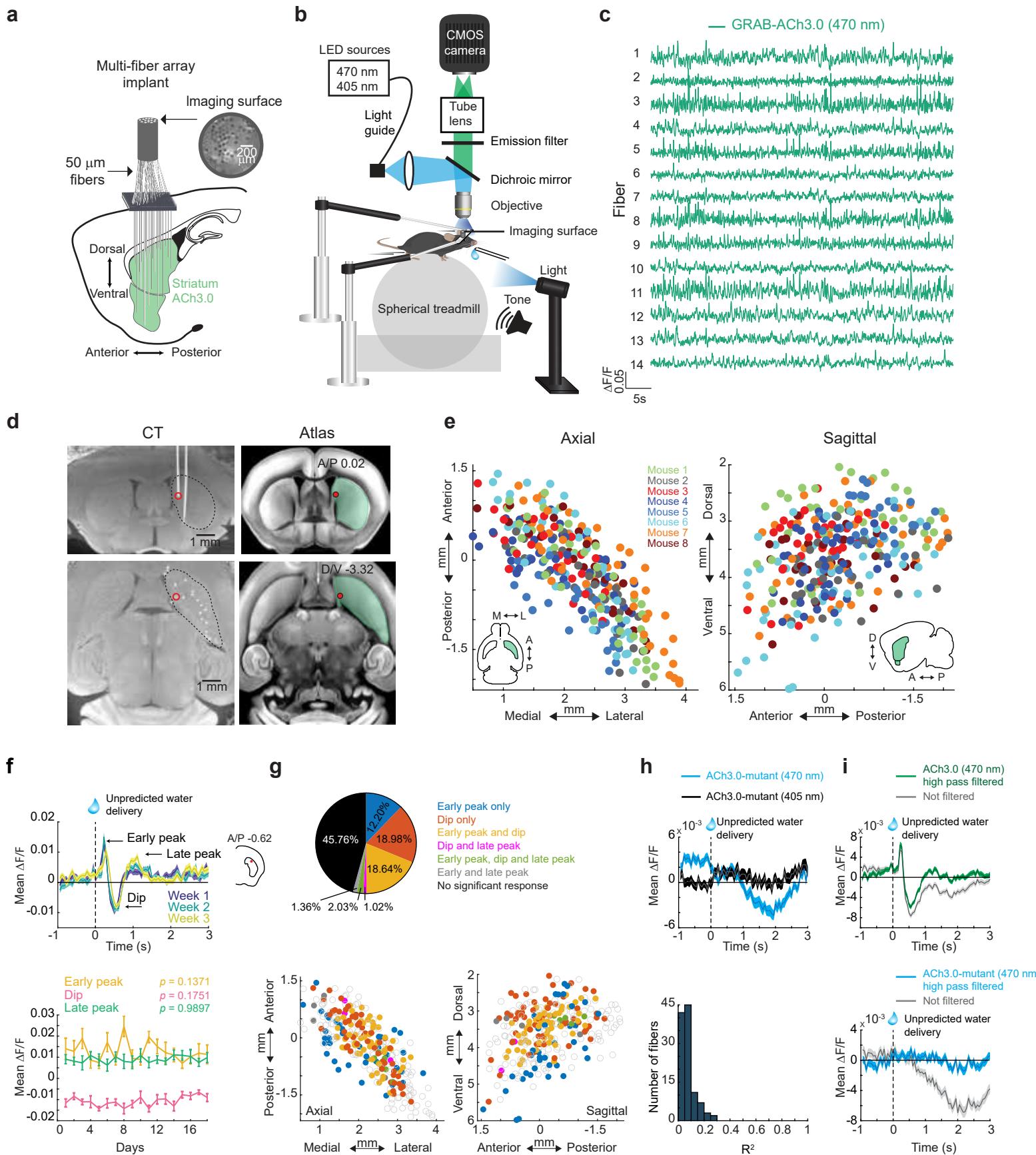
- 1085 32. Kondabolu, K., Roberts, E.A., Bucklin, M., McCarthy, M.M., Kopell, N., and Han, X.  
1086 (2016). Striatal cholinergic interneurons generate beta and gamma oscillations in the  
1087 corticostriatal circuit and produce motor deficits. *Proc Natl Acad Sci U S A* 113, E3159-3168.  
1088 <https://doi.org/10.1073/pnas.1605658113>.
- 1089 33. Huang, Z., Chen, R., Ho, M., Xie, X., Gangal, H., Wang, X., and Wang, J. (2024).  
1090 Dynamic responses of striatal cholinergic interneurons control behavioral flexibility. *Sci. Adv.*  
1091 10, eadn2446. <https://doi.org/10.1126/sciadv.adn2446>.
- 1092 34. Steinberg, E.E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K., and Janak, P.H.  
1093 (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat*  
1094 *Neurosci* 16, 966-973. <https://doi.org/10.1038/nn.3413>.
- 1095 35. Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the  
1096 temporal prediction of reward during learning. *Nat Neurosci* 1, 304-309.  
1097 <https://doi.org/10.1038/1124>.
- 1098 36. Lak, A., Stauffer, W.R., and Schultz, W. (2016). Dopamine neurons learn relative chosen  
1099 value from probabilistic rewards. *eLife* 5, e18044. <https://doi.org/10.7554/eLife.18044>.
- 1100 37. Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C.K., Hassabis, D., Munos, R.,  
1101 and Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement  
1102 learning. *Nature* 577, 671-675. <https://doi.org/10.1038/s41586-019-1924-6>.
- 1103 38. Benhamou, L., Kehat, O., and Cohen, D. (2014). Firing Pattern Characteristics of  
1104 Tonically Active Neurons in Rat Striatum: Context Dependent or Species Divergent? *J.*  
1105 *Neurosci.* 34, 2299-2304. <https://doi.org/10.1523/JNEUROSCI.1798-13.2014>.
- 1106 39. Howe, M., Ridouh, I., Allegra Mascaro, A.L., Larios, A., Azcorra, M., and Dombeck, D.A.  
1107 (2019). Coordination of rapid cholinergic and dopaminergic signaling in striatum during  
1108 spontaneous movement. *eLife* 8, e44903. <https://doi.org/10.7554/eLife.44903>.
- 1109 40. Yarom, O., and Cohen, D. (2011). Putative Cholinergic Interneurons in the Ventral and  
1110 Dorsal Regions of the Striatum Have Distinct Roles in a Two Choice Alternative Association  
1111 Task. *Front. Syst. Neurosci.* 5. <https://doi.org/10.3389/fnsys.2011.00036>.
- 1112 41. Atallah, H.E., McCool, A.D., Howe, M.W., and Graybiel, A.M. (2014). Neurons in the  
1113 ventral striatum exhibit cell-type-specific representations of outcome during learning. *Neuron*  
1114 82, 1145-1156. <https://doi.org/10.1016/j.neuron.2014.04.021>.
- 1115 42. Mohebi, A., Collins, V.L., and Berke, J.D. (2023). Accumbens cholinergic interneurons  
1116 dynamically promote dopamine release and enable motivation. *eLife* 12, e85011.  
1117 <https://doi.org/10.7554/eLife.85011>.
- 1118 43. Marche, K., Martel, A.-C., and Apicella, P. (2017). Differences between Dorsal and  
1119 Ventral Striatum in the Sensitivity of Tonically Active Neurons to Rewarding Events. *Front*  
1120 *Syst Neurosci* 11, 52. <https://doi.org/10.3389/fnsys.2017.00052>.
- 1121 44. Vu, M.-A.T., Brown, E.H., Wen, M.J., Noggle, C.A., Zhang, Z., Monk, K.J., Bouabid, S.,  
1122 Mroz, L., Graham, B.M., Zhuo, Y., et al. (2024). Targeted micro-fiber arrays for measuring and  
1123 manipulating localized multi-scale neural dynamics over large, deep brain volumes during  
1124 behavior. *Neuron* 112, 909-923.e9. <https://doi.org/10.1016/j.neuron.2023.12.011>.
- 1125 45. Jing, M., Li, Y., Zeng, J., Huang, P., Skirzewski, M., Kljakic, O., Peng, W., Qian, T., Tan,  
1126 K., Zou, J., et al. (2020). An optimized acetylcholine sensor for monitoring *in vivo* cholinergic  
1127 activity. *Nat Methods* 17, 1139-1146. <https://doi.org/10.1038/s41592-020-0953-2>.
- 1128 46. Pisanello, M., Pisano, F., Hyun, M., Maglie, E., Balena, A., De Vittorio, M., Sabatini, B.L.,  
1129 and Pisanello, F. (2019). The Three-Dimensional Signal Collection Field for Fiber Photometry  
1130 in Brain Tissue. *Front. Neurosci.* 13, 82. <https://doi.org/10.3389/fnins.2019.00082>.
- 1131 47. Engelbrecht, C.J., Göbel, W., and Helmchen, F. (2009). Enhanced fluorescence signal in  
1132 nonlinear microscopy through supplementary fiber-optic light collection. *Opt. Express*, OE 17,  
1133 6421-6435. <https://doi.org/10.1364/OE.17.006421>.
- 1134 48. Vu, M.-A.T., Brown, E.H., Wen, M.J., Noggle, C.A., Zhang, Z., Monk, K.J., Bouabid, S.,  
1135 Mroz, L., Graham, B.M., Zhuo, Y., et al. (2024). Targeted micro-fiber arrays for measuring and

- 1136 manipulating localized multi-scale neural dynamics over large, deep brain volumes during  
1137 behavior. *Neuron* 112, 909–923.e9. <https://doi.org/10.1016/j.neuron.2023.12.011>.
- 1138 49. Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M.,  
1139 Facer, B., Ho, A., et al. (2020). The Allen Mouse Brain Common Coordinate Framework: A  
1140 3D Reference Atlas. *Cell* 181, 936–953.e20. <https://doi.org/10.1016/j.cell.2020.04.007>.
- 1141 50. Krok, A.C., Maltese, M., Mistry, P., Miao, X., Li, Y., and Tritsch, N.X. (2023). Intrinsic  
1142 dopamine and acetylcholine dynamics in the striatum of mice. *Nature* 621, 543–549.  
1143 <https://doi.org/10.1038/s41586-023-05995-9>.
- 1144 51. Zhang, W.-T., Chao, T.-H.H., Yang, Y., Wang, T.-W., Lee, S.-H., Oyarzabal, E.A., Zhou,  
1145 J., Nonneman, R., Pegard, N.C., Zhu, H., et al. (2022). Spectral fiber photometry derives  
1146 hemoglobin concentration changes for accurate measurement of fluorescent sensor activity.  
1147 *Cell Reports Methods* 2, 100243. <https://doi.org/10.1016/j.crmeth.2022.100243>.
- 1148 52. Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L., and Tank, D.W. (2010). Functional  
1149 imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat  
1150 Neurosci* 13, 1433–1440. <https://doi.org/10.1038/nn.2648>.
- 1151 53. Hunnicutt, B.J., Jongbloets, B.C., Birdsong, W.T., Gertz, K.J., Zhong, H., and Mao, T.  
1152 (2016). A comprehensive excitatory input map of the striatum reveals novel functional  
1153 organization. *eLife* 5, e19103. <https://doi.org/10.7554/eLife.19103>.
- 1154 54. Hintiryan, H., Foster, N.N., Bowman, I., Bay, M., Song, M.Y., Gou, L., Yamashita, S.,  
1155 Bienkowski, M.S., Zingg, B., Zhu, M., et al. (2016). The mouse cortico-striatal projectome.  
1156 *Nat Neurosci* 19, 1100–1114. <https://doi.org/10.1038/nn.4332>.
- 1157 55. Schultz, W., Dayan, P., and Montague, P.R. (1997). A Neural Substrate of Prediction and  
1158 Reward. *Science* 275, 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>.
- 1159 56. van Elzelingen, W., Goedhoop, J., Warnaar, P., Denys, D., Arbab, T., and Willuhn, I.  
1160 (2022). A unidirectional but not uniform striatal landscape of dopamine signaling for  
1161 motivational stimuli. *Proceedings of the National Academy of Sciences* 119, e2117270119.  
1162 <https://doi.org/10.1073/pnas.2117270119>.
- 1163 57. Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning  
1164 mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci* 10,  
1165 1020–1028. <https://doi.org/10.1038/nn1923>.
- 1166 58. Gritton, H.J., Howe, W.M., Romano, M.F., DiFeliceantonio, A.G., Kramer, M.A.,  
1167 Saligrama, V., Bucklin, M.E., Zemel, D., and Han, X. (2019). Unique contributions of  
1168 parvalbumin and cholinergic interneurons in organizing striatal networks during movement.  
1169 *Nat Neurosci* 22, 586–597. <https://doi.org/10.1038/s41593-019-0341-3>.
- 1170 59. Shroff, S.N., Lowet, E., Sridhar, S., Gritton, H.J., Abumuaileq, M., Tseng, H.-A., Cheung,  
1171 C., Zhou, S.L., Kondabolu, K., and Han, X. (2023). Striatal cholinergic interneuron membrane  
1172 voltage tracks locomotor rhythms in mice. *Nat Commun* 14, 3802.  
1173 <https://doi.org/10.1038/s41467-023-39497-z>.
- 1174 60. Chantranupong, L., Beron, C.C., Zimmer, J.A., Wen, M.J., Wang, W., and Sabatini, B.L.  
1175 (2023). Dopamine and glutamate regulate striatal acetylcholine in decision-making. *Nature*  
1176 621, 577–585. <https://doi.org/10.1038/s41586-023-06492-9>.
- 1177 61. Kim, T., Capps, R.A., Hamade, K.C., Barnett, W.H., Todorov, D.I., Latash, E.M., Markin,  
1178 S.N., Rybak, I.A., and Molkov, Y.I. (2019). The Functional Role of Striatal Cholinergic  
1179 Interneurons in Reinforcement Learning From Computational Perspective. *Front. Neural  
1180 Circuits* 13. <https://doi.org/10.3389/fncir.2019.00010>.
- 1181 62. Deffains, M., and Bergman, H. (2015). Striatal cholinergic interneurons and  
1182 cortico-striatal synaptic plasticity in health and disease. *Movement Disorders* 30, 1014–1025.  
1183 <https://doi.org/10.1002/mds.26300>.
- 1184 63. Zhang, Y.-F., and Cragg, S.J. (2017). Pauses in Striatal Cholinergic Interneurons: What  
1185 is Revealed by Their Common Themes and Variations? *Front. Syst. Neurosci.* 11.  
1186 <https://doi.org/10.3389/fnsys.2017.00080>.

- 1187 64. Pan, W.-X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2008). Tripartite mechanism of  
1188 extinction suggested by dopamine neuron activity and temporal difference model. *J Neurosci*  
1189 28, 9619–9631. <https://doi.org/10.1523/JNEUROSCI.0255-08.2008>.
- 1190 65. Chuhma, N., Mingote, S., Moore, H., and Rayport, S. (2014). Dopamine Neurons Control  
1191 Striatal Cholinergic Neurons via Regionally Heterogeneous Dopamine and Glutamate  
1192 Signaling. *Neuron* 81, 901–912. <https://doi.org/10.1016/j.neuron.2013.12.027>.
- 1193 66. Straub, C., Tritsch, N.X., Hagan, N.A., Gu, C., and Sabatini, B.L. (2014). Multiphasic  
1194 Modulation of Cholinergic Interneurons by Nigrostriatal Afferents. *Journal of Neuroscience*  
1195 34, 8557–8569. <https://doi.org/10.1523/JNEUROSCI.0589-14.2014>.
- 1196 67. Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.-H., Folk, R.W.,  
1197 Broussard, G.J., Liang, R., Jang, M.J., et al. (2018). Ultrafast neuronal imaging of dopamine  
1198 dynamics with designed genetically encoded sensors. *Science* 360, eaat4422.  
1199 <https://doi.org/10.1126/science.aat4422>.
- 1200 68. Marvin, J.S., Scholl, B., Wilson, D.E., Podgorski, K., Kazemipour, A., Müller, J.A.,  
1201 Schoch, S., Quiroz, F.J.U., Rebola, N., Bao, H., et al. (2018). Stability, affinity, and chromatic  
1202 variants of the glutamate sensor iGluSnFR. *Nat Methods* 15, 936–939.  
1203 <https://doi.org/10.1038/s41592-018-0171-3>.
- 1204 69. Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and Bergman, H. (2008). Midbrain  
1205 Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between  
1206 Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning  
1207 Trials. *J. Neurosci.* 28, 11673–11684. <https://doi.org/10.1523/JNEUROSCI.3839-08.2008>.
- 1208 70. Azcorra, M., Gaertner, Z., Davidson, C., He, Q., Kim, H., Nagappan, S., Hayes, C.K.,  
1209 Ramakrishnan, C., Fenno, L., Kim, Y.S., et al. (2023). Unique functional responses  
1210 differentially map onto genetic subtypes of dopamine neurons. *Nat Neurosci* 26, 1762–1774.  
1211 <https://doi.org/10.1038/s41593-023-01401-9>.
- 1212 71. Lerner, T.N., Shilyansky, C., Davidson, T.J., Evans, K.E., Beier, K.T., Zalocusky, K.A.,  
1213 Crow, A.K., Malenka, R.C., Luo, L., Tomer, R., et al. (2015). Intact-Brain Analyses Reveal  
1214 Distinct Information Carried by SNC Dopamine Subcircuits. *Cell* 162, 635–647.  
1215 <https://doi.org/10.1016/j.cell.2015.07.014>.
- 1216 72. Howe, M.W., and Dombeck, D.A. (2016). Rapid signalling in distinct dopaminergic axons  
1217 during locomotion and reward. *Nature* 535, 505–510. <https://doi.org/10.1038/nature18942>.
- 1218 73. Tsutsui-Kimura, I., Matsumoto, H., Akiti, K., Yamada, M.M., Uchida, N., and  
1219 Watabe-Uchida, M. (2020). Distinct temporal difference error signals in dopamine axons in  
1220 three regions of the striatum in a decision-making task. *eLife* 9, e62390.  
1221 <https://doi.org/10.7554/eLife.62390>.
- 1222 74. Reynolds, J.N., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of  
1223 reward-related learning. *Nature* 413, 67–70. <https://doi.org/10.1038/35092560>.
- 1224 75. Wang, Z., Kai, L., Day, M., Ronesi, J., Yin, H.H., Ding, J., Tkatch, T., Lovinger, D.M., and  
1225 Surmeier, D.J. (2006). Dopaminergic control of corticostriatal long-term synaptic depression  
1226 in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* 50, 443–452.  
1227 <https://doi.org/10.1016/j.neuron.2006.04.010>.
- 1228 76. Delamater, A.R., and Westbrook, R.F. (2014). Psychological and neural mechanisms of  
1229 experimental extinction: A selective review. *Neurobiology of Learning and Memory* 108,  
1230 38–51. <https://doi.org/10.1016/j.nlm.2013.09.016>.
- 1231 77. Shen, W., Plotkin, J.L., Francardo, V., Ko, W.K.D., Xie, Z., Li, Q., Fieblinger, T., Wess, J.,  
1232 Neubig, R.R., Lindsley, C.W., et al. (2015). M4 Muscarinic Receptor Signaling Ameliorates  
1233 Striatal Plasticity Deficits in Models of L-DOPA-Induced Dyskinesia. *Neuron* 88, 762–773.  
1234 <https://doi.org/10.1016/j.neuron.2015.10.039>.
- 1235 78. Nair, A.G., Castro, L.R.V., El Khoury, M., Gorgievski, V., Giros, B., Tzavara, E.T.,  
1236 Hellgren-Kotaleski, J., and Vincent, P. (2019). The high efficacy of muscarinic M4 receptor in  
1237 D1 medium spiny neurons reverses striatal hyperdopaminergia. *Neuropharmacology* 146,

- 1238 74–83. <https://doi.org/10.1016/j.neuropharm.2018.11.029>.
- 1239 79. Lee, S.J., Lodder, B., Chen, Y., Patriarchi, T., Tian, L., and Sabatini, B.L. (2021).
- 1240 Cell-type-specific asynchronous modulation of PKA by dopamine in learning. *Nature* 590,
- 1241 451–456. <https://doi.org/10.1038/s41586-020-03050-5>.
- 1242 80. Shen, W., Hamilton, S.E., Nathanson, N.M., and Surmeier, D.J. (2005). Cholinergic
- 1243 Suppression of KCNQ Channel Currents Enhances Excitability of Striatal Medium Spiny
- 1244 Neurons. *J. Neurosci.* 25, 7449–7458. <https://doi.org/10.1523/JNEUROSCI.1381-05.2005>.
- 1245 81. Pisanello, M., Pisano, F., Hyun, M., Maglie, E., Balena, A., De Vittorio, M., Sabatini, B.L.,
- 1246 and Pisanello, F. (2019). The Three-Dimensional Signal Collection Field for Fiber Photometry
- 1247 in Brain Tissue. *Front. Neurosci.* 13. <https://doi.org/10.3389/fnins.2019.00082>.
- 1248 82. Perkins, L.N., Semu, D., Shen, J., Boas, D.A., and Gardner, T.J. (2018). High-density
- 1249 microfibers as a potential optical interface to reach deep brain regions. *J. Neural Eng.* 15,
- 1250 066002. <https://doi.org/10.1088/1741-2552/aadbb2>.
- 1251 83. Sergejeva, M., Papp, E.A., Bakker, R., Gaudnek, M.A., Okamura-Oho, Y., Boline, J.,
- 1252 Bjaalie, J.G., and Hess, A. (2015). Anatomical landmarks for registration of experimental
- 1253 image data to volumetric rodent brain atlasing templates. *J Neurosci Methods* 240, 161–169.
- 1254 <https://doi.org/10.1016/j.jneumeth.2014.11.005>.
- 1255 84. Zhang, Y., Zhao, S., Rodriguez, E., Takatoh, J., Han, B.-X., Zhou, X., and Wang, F.
- 1256 (2015). Identifying local and descending inputs for primary sensory neurons. *J Clin Invest*
- 1257 125, 3782–3794. <https://doi.org/10.1172/JCI81156>.
- 1258 85. Miri, A., Daie, K., Arrenberg, A.B., Baier, H., Aksay, E., and Tank, D.W. (2011). Spatial
- 1259 gradients and multidimensional dynamics in a neural integrator circuit. *Nat Neurosci* 14,
- 1260 1150–1159. <https://doi.org/10.1038/nn.2888>.

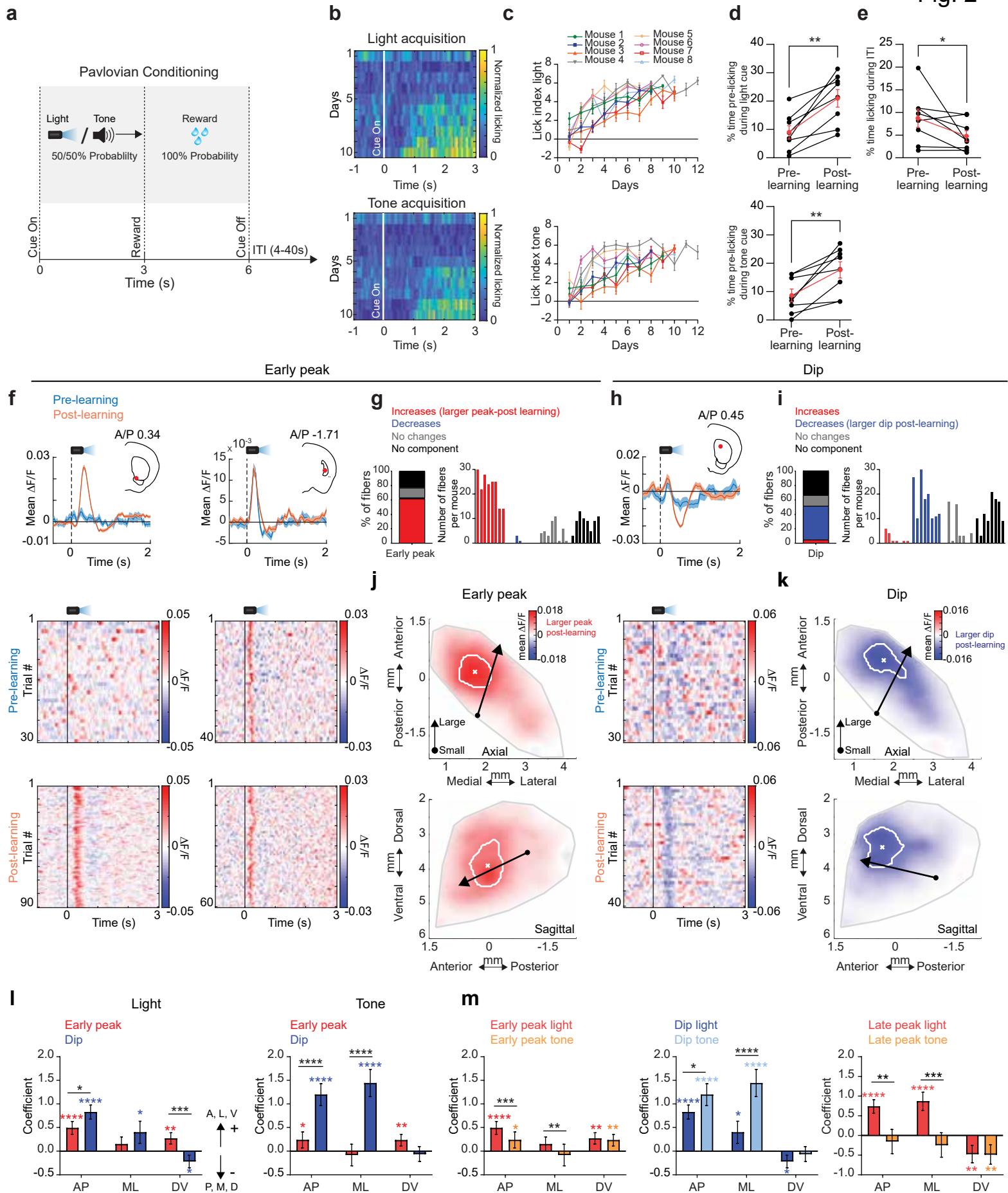
Fig. 1



1261 **FIGURE TITLES AND LEGENDS**

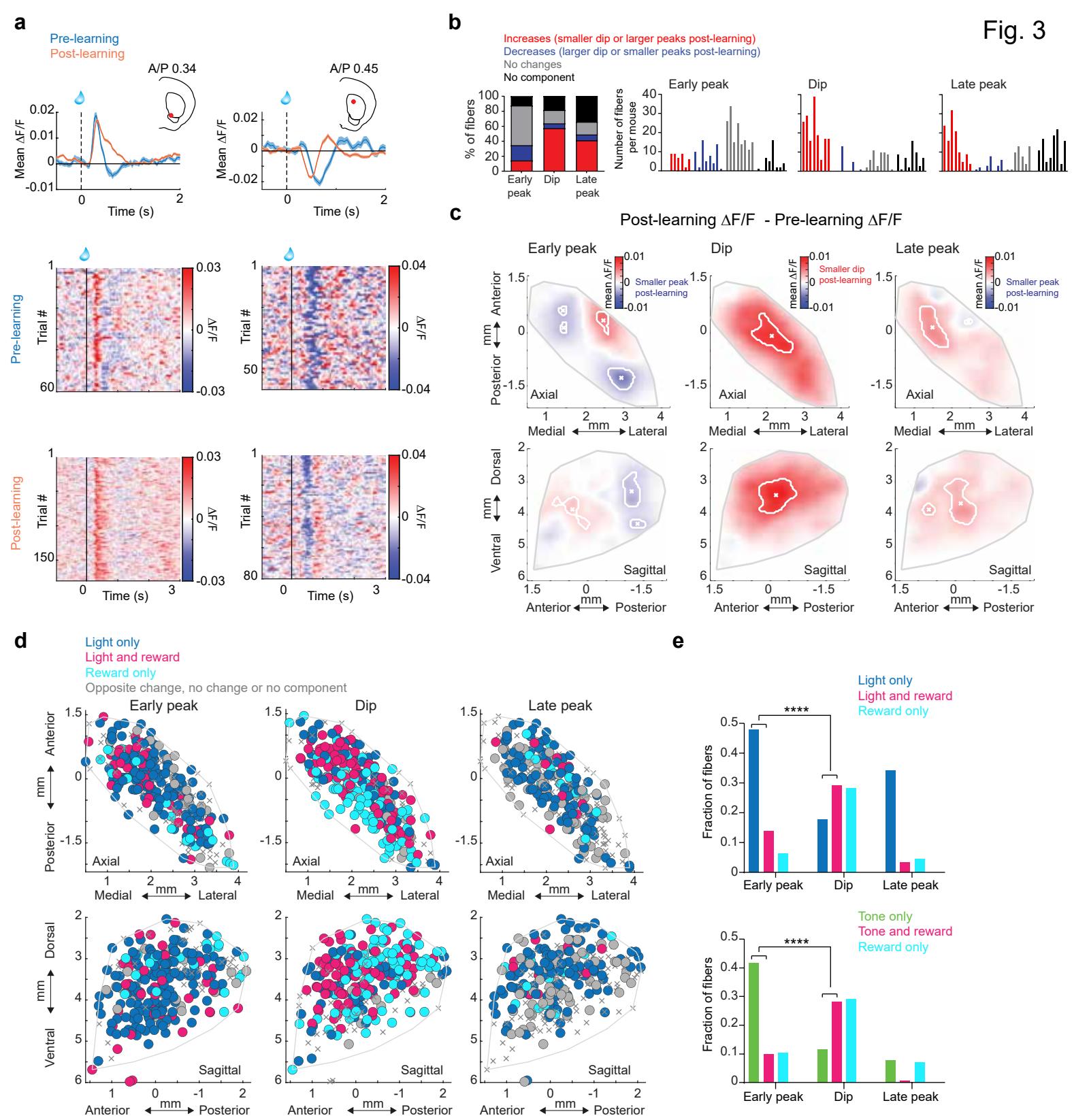
1262 **Fig. 1: Chronic striatum-wide measurements of rapid ACh release dynamics in**  
1263 **head-fixed, behaving mice.** **a** Schematic of the fiber array approach for measuring ACh  
1264 release across the striatum. **b** Spherical treadmill and imaging set up for head-fixed mice. **c**  
1265 ACh3.0 fluorescence traces ( $\Delta F/F$ ) measured from 14 example fibers in the striatum of a single  
1266 mouse. **d** Left: Post-mortem micro-CT scan images in the coronal plane (top) and axial plane  
1267 (bottom) from a representative mouse. Fibers appear in white. Right: Images from the Allen  
1268 Brain Common Coordinate Framework Atlas corresponding to the CT planes on left. Red circles  
1269 indicate the position of an automatically localized fiber tip. **e** Locations of all fibers used for ACh  
1270 measurements in the axial (left) and sagittal (right) planes. Each dot is the location (relative to  
1271 bregma) of a single fiber and the color indicates the mouse identity ( $n = 8$  mice). **f** Top: mean  
1272  $\Delta F/F$  aligned to unpredicted water reward delivery for a single fiber for single sessions ( $n=8$   
1273 reward deliveries/session) at 1-3 weeks from the first day of Pavlovian training (Fig. 2). Inset  
1274 indicates fiber location in the coronal plane. Shaded regions, S.E.M. Bottom: The mean peak  
1275 (early/late) or dip  $\Delta F/F$  for unpredicted reward deliveries across 18 consecutive sessions of  
1276 training for the example at top. Values are the mean  $\pm$  S.E.M. For each component, the mean of  
1277 each day was compared to the mean of every other day using one-way ANOVA. **g** Top:  
1278 percentage of the total fibers (295 fibers) across all mice with significant ( $p < 0.01$ , two-tailed  
1279 Wilcoxon rank-sum test, see Methods) component(s) to unpredicted reward delivery. Bottom:  
1280 maps as in (e) showing the presence of each combination of signal components to unpredicted  
1281 reward delivery for each fiber. Empty circles indicate no significant response. **h** Top: Mean  $\Delta F/F$   
1282 aligned to unpredicted reward delivery with quasi-simultaneous 405nm (black) and 470nm  
1283 (blue) illumination in a mouse expressing the non-functional mutant ACh3.0 sensor. Shaded  
1284 regions, S.E.M. Bottom: Histogram of Pearson's correlation coefficients between  $\Delta F/F$   
1285 fluorescence traces obtained with quasi-simultaneous 405nm and 470nm illumination in mutant  
1286 ACh sensor expressing mice ( $n = 3$ ). **i** Mean unpredicted reward triggered  $\Delta F/F$  from a single  
1287 fiber in functional ACh3.0 sensor (top) and mutant ACh sensor (bottom) expressing mice before  
1288 and after 0.3Hz high pass filtering. Shaded region, S.E.M. Note that the small, slow artifactual  
1289 decrease in the mutant sensor recording is largely eliminated with filtering, but the rapid reward  
1290 triggered release measured with the functional sensor is preserved.

Fig. 2



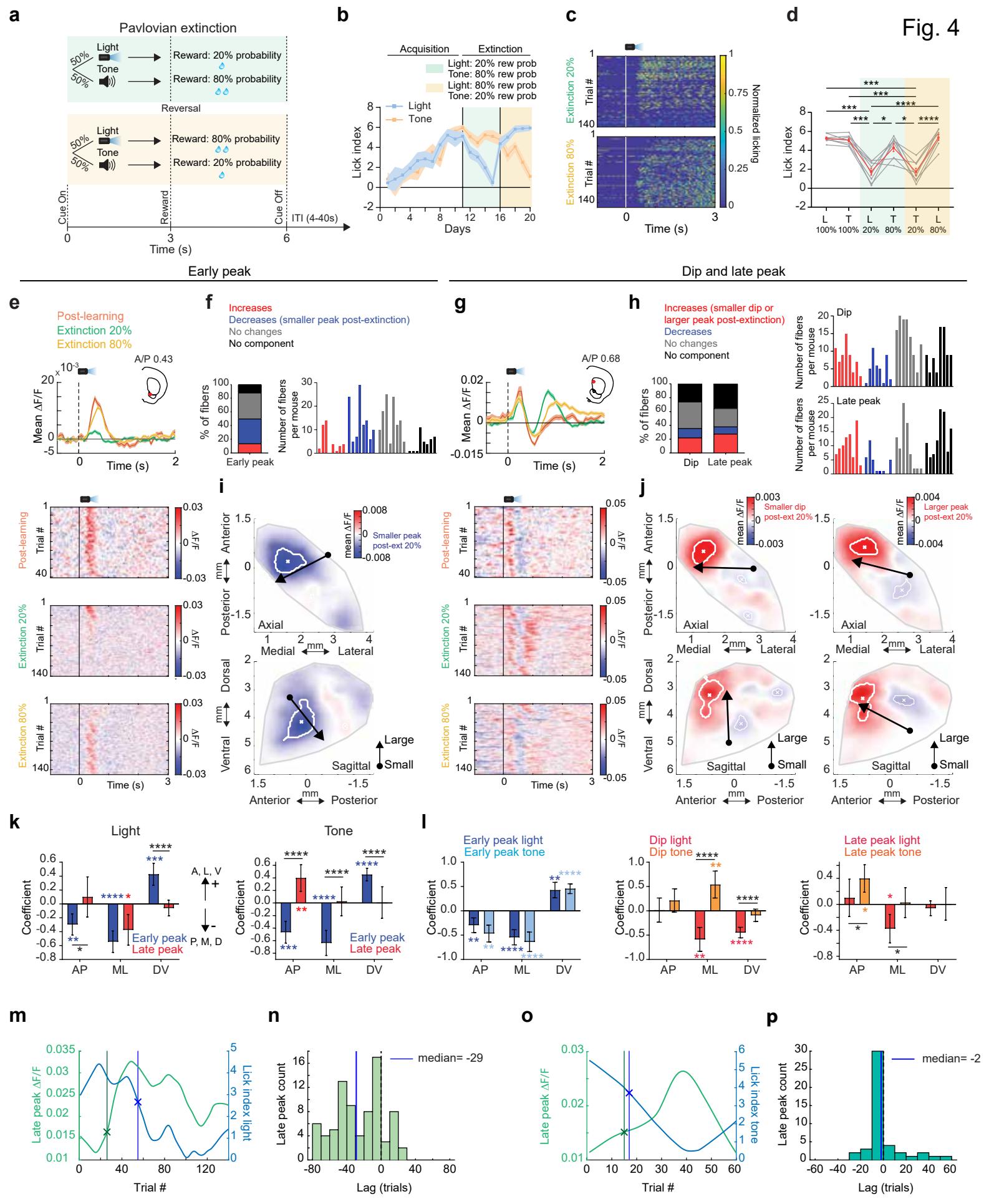
**1291 Fig. 2: ACh release components for conditioned cues evolve bi-directionally over distinct**  
**1292 anatomical gradients during Pavlovian learning.** **a** Schematic of dual-cue delay Pavlovian  
1293 conditioning task. **b** Normalized mean lick spout contacts aligned to light (top) and tone (bottom)  
1294 cue onset for 10 consecutive sessions from the start of training for one representative mouse. **c**  
1295 Mean lick index on light (top) and tone (bottom) cue trials for each day after the start of  
1296 Pavlovian training for each mouse. Error bars, S.E.M. **d** Mean percent time spent licking during  
1297 the 3s light (top) and tone (bottom) cue interval prior to reward for each mouse for sessions pre  
1298 and post learning. Red points are means ( $\pm$  S.E.M.) across all mice ( $n = 8$ ). \*\* $p < 0.01$ , two-tailed  
1299 Wilcoxon matched-pairs signed rank test. **e** Same as (d) for licking during ITI periods (\* $p < 0.05$ ).  
1300 **f** Top: Mean  $\Delta F/F$  aligned to the light cue onset for 2 representative fibers for trials pre (blue)  
1301 and post (orange) learning. Left example, significant early peak change; right, no change.  
1302 Shaded regions, S.E.M. Red dots in insets indicate the fiber locations in the coronal plane.  
1303 Bottom: Light-cue-aligned  $\Delta F/F$  for all trials included in the triggered averages at top. **g** Left:  
1304 percent of all fibers with significant increases or decreases, no change, or no significant  
1305 component from pre to post learning for the early peak  $\Delta F/F$  at light cue onset. Right: histogram  
1306 of number of fibers per mouse with significant early peak changes. Each bar is the fiber count  
1307 for one mouse for each condition indicated by colors at left. **h** Mean  $\Delta F/F$  aligned to the light cue  
1308 onset as in (f) for a representative fiber with a significant change in the dip with learning. **i** Same  
1309 as (g) for dip  $\Delta F/F$  changes from pre to post learning. **j** Maps (axial, top; sagittal, bottom)  
1310 showing spatially weighted means across locations of differences with learning (post - pre  $\Delta F/F$ )  
1311 for the mean early peak  $\Delta F/F$  at light cue onset. Lines indicate the axes of maximal variation  
1312 and arrows indicate the direction of peak increases from smallest to largest changes. White  
1313 contours indicate regions with changes in the highest 10<sup>th</sup> percentile. **k** Same as (j) for the dip  
1314 component. **l** Model coefficients indicating the relative magnitude and direction of the variation in  
1315 mean  $\Delta F/F$  differences with learning for the peak and dip components for light and tone cues  
1316 across each striatal axis (AP: anterior-posterior, ML: medial-lateral, DV: dorsal-ventral). The sign  
1317 of the coefficient indicates the direction of the largest differences (see arrows). Error bars,  
1318 S.E.M. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , Wald t-test on model coefficients. Significant  
1319 interaction terms (black) indicate difference in coefficients between peak and dip for a given  
1320 axis. **m** Same as (l) but comparing spatial coefficients for tone and light for all three  
1321 components.

Fig. 3



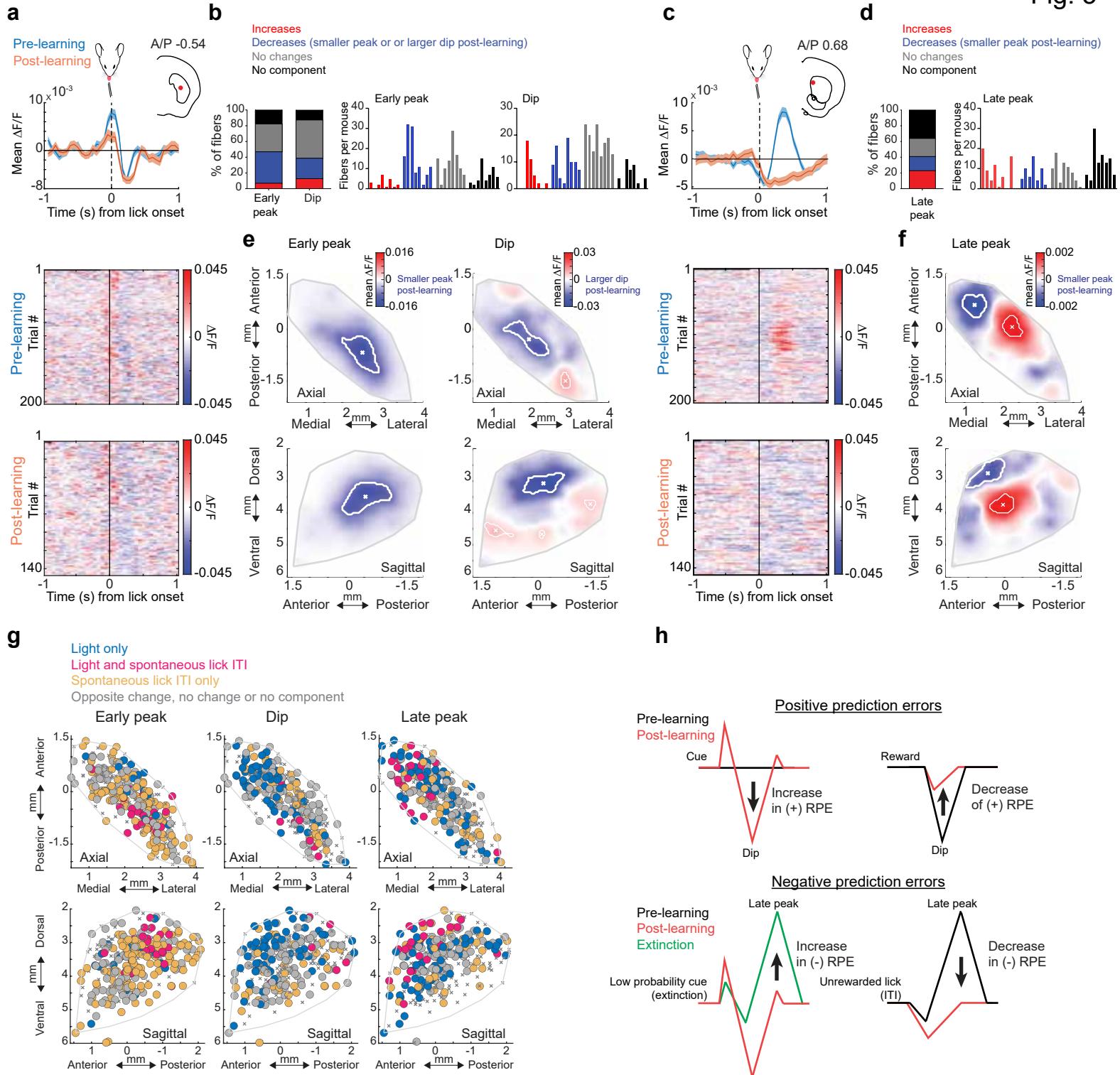
**Fig. 3: Opposite ACh changes with learning at cue and reward are signal component selective and spatially concentrated in the anterior dorsal striatum.** **a** Top: Mean  $\Delta F/F$  aligned to the reward delivery in the task for 2 representative fibers (same fibers as Fig. 2f, left and 2h) for trials pre (blue) and post (orange) learning. Significant changes for both are present in the dip, but not the early peak. Shaded regions, S.E.M. Red dots in insets indicate the fiber locations in the coronal plane. Bottom: Reward-triggered  $\Delta F/F$  for all trials included in the triggered averages at top. **b** Left: percent of all fibers with significant increases or decreases, no change, or no significant component from pre to post learning for each signal component  $\Delta F/F$  at reward consumption onset. Right: histograms of the number of fibers per mouse with significant changes for each component. Each bar is the fiber count for one mouse for each condition indicated by colors at left. **c** Maps (axial, top; sagittal, bottom) showing spatially weighted means across locations of differences with learning (post - pre) for the mean  $\Delta F/F$  for the three signal components at reward consumption onset. White contours indicate regions with changes in the highest 10<sup>th</sup> percentile. **d** Maps showing each fiber (dot) color coded according to whether significant changes from pre to post learning (see Methods) were present at the light cue onset only (dark blue), reward consumption only (light blue) or both (pink). Pink dots indicate locations where the component magnitude became larger (dip more negative or peak more positive) for cue and smaller for reward over learning, consistent with reward prediction error encoding. **e** Fraction of all fibers for each component classified according to changes across learning for light cue (top) or tone cue (bottom) and reward as indicated in (d) ( $p=8.1\times 10^{-13}$  early peak vs. dip between light only and light and reward,  $p=7.8\times 10^{-18}$  early peak vs dip between tone only and tone and reward, \*\*\* $p < 0.0001$ , Fisher's exact test).

Fig. 4



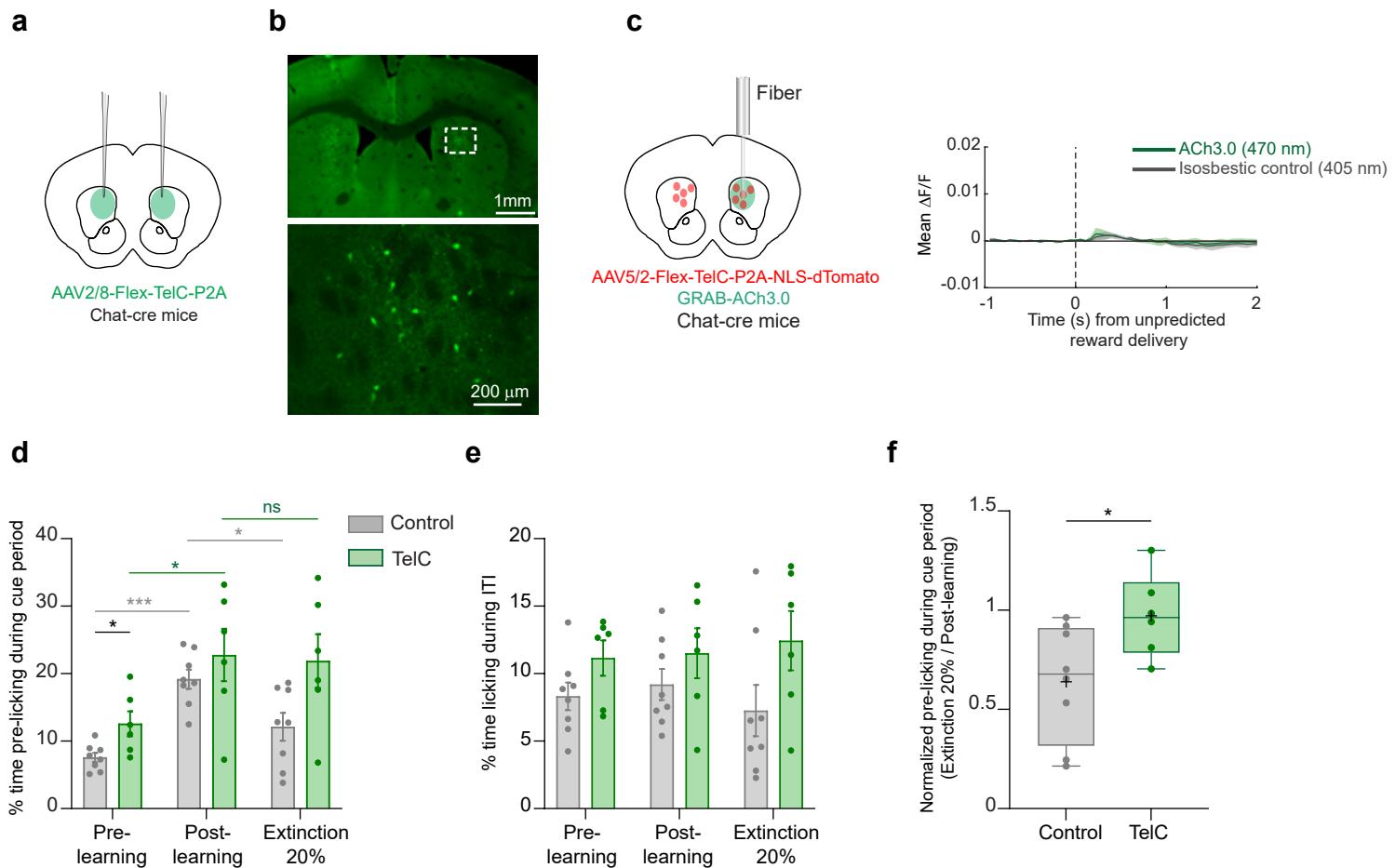
**Figure 4: Emergence of cue-evoked ACh increases in the anterior dorsal striatum during partial extinction.** **a** Schematic of the partial extinction paradigm. **b** Mean lick index on light and tone cue trials for each day of training for one representative mouse. Shaded regions, S.E.M. across trials for each day. **c** Rasters of normalized licking aligned to light cue onset for all trials from the beginning of extinction training for a single mouse. **d** Lick indices across all mice ( $n = 8$ ) for the light and tone cues associated with different reward probabilities during initial learning (100%) and partial extinction (20% and 80%). Mean across mice shown in red, individual mice in gray. Error bars, S.E.M. Repeated measures one-way ANOVA, followed by Tukey *post hoc* test. \*:  $p < 0.05$ , \*\*:  $p < 0.001$ , \*\*\*\*:  $p < 0.0001$ . T: tone, L: light. **e** Top: Mean  $\Delta F/F$  aligned to the light cue onset for a representative fiber with a significant decrease in the early peak from learning to extinction for trials where light was associated with 100% (post-learning), 80%, and 20% (extinction) probabilities. Shaded regions, S.E.M. Red dots in insets indicate the fiber locations in the coronal plane. Bottom: Light-triggered  $\Delta F/F$  for all trials included in the triggered averages at top. **f** Left: percent of all fibers with significant increases or decreases, no change, or no significant component from post learning to the 20% light extinction phase for the early peak  $\Delta F/F$  at light cue onset. Right: histogram of number of fibers per mouse with significant early peak changes. Each bar is the fiber count for one mouse for each condition indicated by colors at left. **g** Mean  $\Delta F/F$  aligned to the light cue onset as in (e) for a representative fiber with a significant increase in the late peak and decrease in the dip from late learning to the extinction phases. **h** Left: percent of all fibers with changes from post learning to extinction as in (f) for the dip and late peak components. Right: histogram of number of fibers per mouse with significant dip (top) and late peak (bottom) changes. Each bar is the fiber count for one mouse for each condition indicated by colors at left. **i** Maps (axial, top; sagittal, bottom) showing spatially weighted means across locations of differences between the 20% extinction and 100% post learning phases ( $20\% - 100\% \Delta F/F$ ) for the mean early peak  $\Delta F/F$  at light cue onset. Lines indicate the axes of maximal variation and arrows indicate the direction of peak decreases from smallest to largest changes. White contours indicate regions with changes in the highest 10<sup>th</sup> percentile. **j** Same as (i) but for dip (left) and late peak (right). **k** Model coefficients indicating the relative magnitude and direction variation of the mean  $\Delta F/F$  differences between the 20% extinction and 100% post learning phases for the early and late peak components for light and tone cues across each striatal axis (AP: anterior-posterior, ML: medial-lateral, DV: dorsal-ventral). The sign of the coefficient indicates the direction of the largest differences (see arrows). Error bars, S.E.M. \* $p < 0.05$ , \*\* $p < 0.01$  \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , Wald t-test followed by bonferroni *post hoc* analysis on model coefficients. Significant interaction terms (black) indicate difference in coefficients between early and late peak for a given axis. **l** Same as (k) but comparing spatial coefficients for tone and light for all three components. **m** Mean light cue-evoked late peak  $\Delta F/F$  (green) and lick index (blue) for all trials following the transition from 100% reward probability to 20% reward probability for a single fiber. The lines indicate where each measure significantly (CUSUM algorithm, see Methods) changed relative to the 100% probability phase. **n** Histogram showing the # fibers with relative latencies between the significant increase in light cue-evoked late peak  $\Delta F/F$  and the decrease in lick index following high to low reward probability transitions. Vertical line indicates the median of the distribution. **o** Same as (m) but for tone. **p** Same as (n), but for tone.

Fig. 5



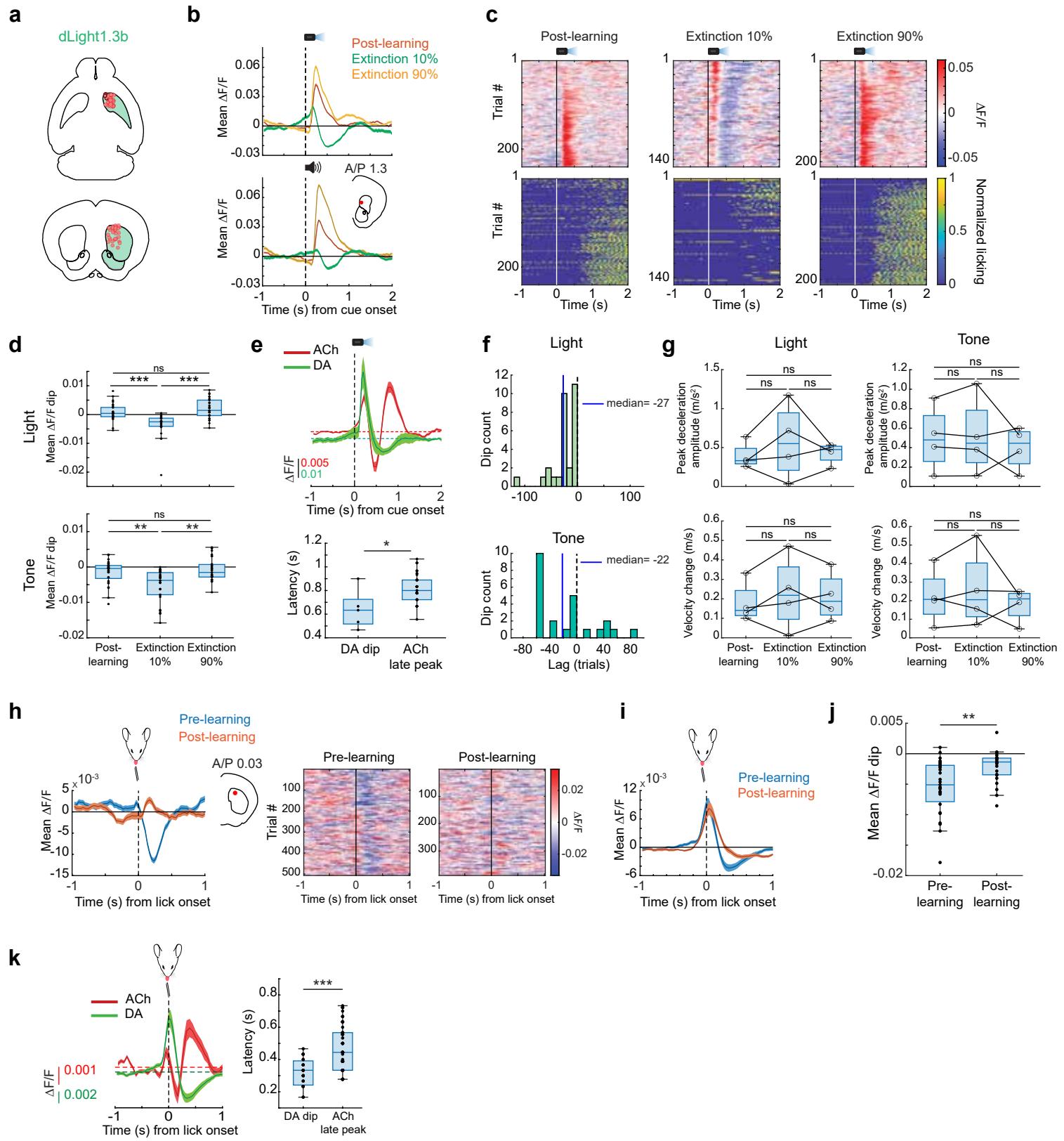
**Fig. 5: Regional changes in ACh release components to unrewarded ITI licks during learning.** **a** Top: Mean  $\Delta F/F$  aligned to unrewarded ITI lick bout onsets for one representative fiber with significant early peak and dip changes for trials pre (blue) and post (orange) Pavlovian learning. Shaded regions, S.E.M. Red dot in inset indicates the fiber location in the coronal plane. Bottom: Lick onset-aligned  $\Delta F/F$  for all trials included in the triggered averages at top. **b** Left: percent of all fibers with significant increases or decreases, no change, or no significant component from pre to post learning for the early peak and dip  $\Delta F/F$  at ITI lick bout onsets. Right: histogram of number of fibers per mouse with significant early peak and dip changes. Each bar is the fiber count for one mouse for each condition indicated by colors at left. **c** Same as **(a)** for a fiber with a significant late peak change. **d** Same as **(b)** for late peak changes with learning. **e** Maps (axial, top; sagittal, bottom) showing spatially weighted means across locations (see Methods) of differences with learning (post - pre  $\Delta F/F$ ) for the mean early peak (left) and dip (right)  $\Delta F/F$  at ITI lick bout onsets. White contours indicate regions with changes in the highest 10<sup>th</sup> percentile. **f** Same as **(e)** for late peak. **g** Maps (axial, top; sagittal, bottom) showing each fiber (dot) color coded according to whether significant changes from pre to post learning (see Methods) were present at the light cue onset only (dark blue), unrewarded ITI lick only (light blue) or both (pink). Pink dots indicate locations where the component magnitude was larger with extinction (dip more negative or peak more positive) for cue and smaller post learning for the unrewarded ITI lick, consistent with negative reward prediction error encoding. **h** Schematic summarizing the changes in ACh release in the aDS region with learning and extinction consistent with positive and negative RPE encoding for distinct signal components.

Fig. 6



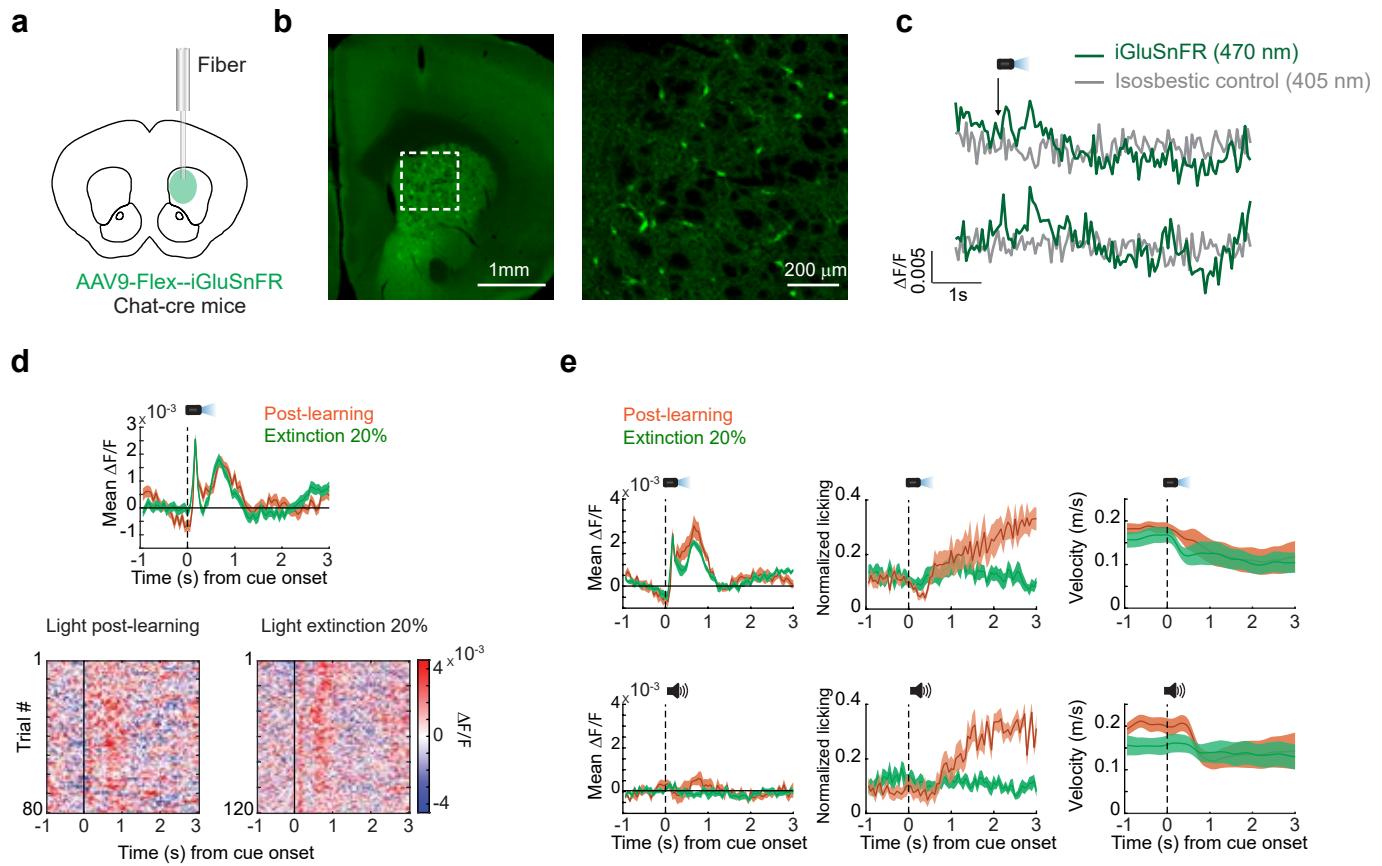
**Fig. 6: Inactivation of ACh release in the anterior dorsal striatum impairs behavioral changes during extinction.** **a** Schematic showing the bilateral injection strategy for the selective expression of TelC in CINs in aDS. **b** Left: Fluorescence image in a coronal section of the striatum showing bilateral TelC expression in aDS for a representative mouse. Right: Magnification of the boxed area in left image showing selective expression of TelC in CINs. **c** Left: Schematic showing the strategy for verifying effects of TelC mediated suppression of CINs on ACh release to unpredicted rewards. Right: Mean  $\Delta F/F$  across all trials ( $n= 209$ ) aligned to unpredicted reward delivery with quasi-simultaneous 405nm (gray) and 470nm (green) illumination in two mice expressing the ACh3.0 sensor and TeLC in CINs. Shaded regions, S.E.M. **d** Mean ( $\pm$  S.E.M.) % time spent licking during the 3s of cue presentation (both light and tone cues merged) for pre learning, post learning and extinction (20% reward probability cue only) for both TelC and control mice. Each dot is an individual mouse ( $n=8$  for control and  $n=6$  for TelC mice). \* $p<0.05$ , \*\*\* $p<0.001$ , two-way ANOVA followed by Tukey *post hoc* analysis for comparisons within and between TelC and control mice. **e** same as (c) for licking during ITI periods. **f** Box plots showing the ratio of mean cue period licking during partial extinction (20% reward probability, light and tone combined) relative to post learning (100% reward probability), indicating no change in cue licking in TelC mice and a reduction in normalized cue licking in control mice during partial extinction. Each dot represents an individual mouse ( $n=8$  for control and  $n=6$  for TelC mice). \* $p<0.05$ , Mann-Whitney test (two-tailed).

Fig. 7



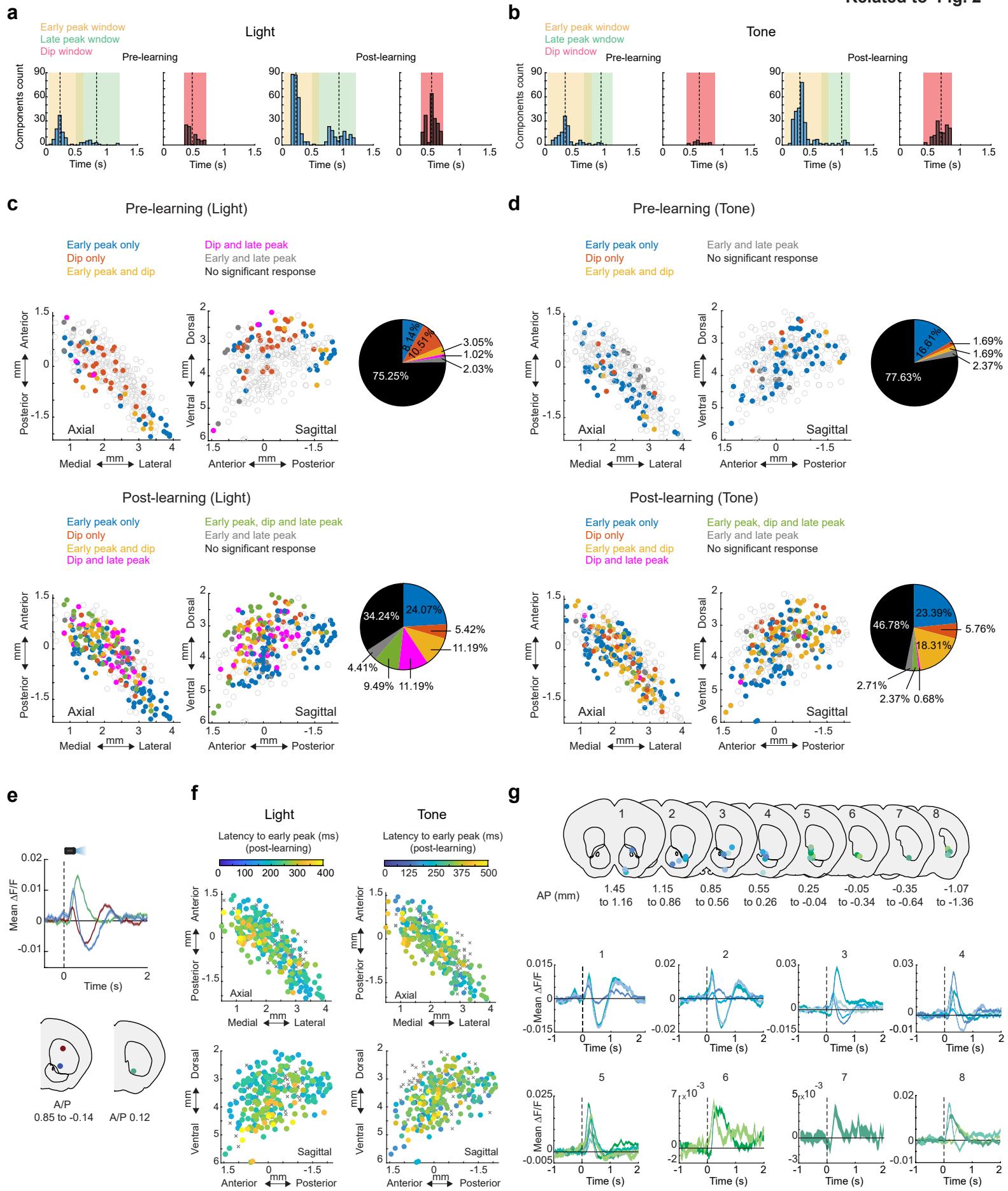
**Fig. 7: Dips in dopamine release in the anterior dorsal striatum are present for unrewarded lick onsets before learning and for cues after extinction.** **a** Schematic showing the locations of fibers ( $n = 29$  fibers in 4 mice) in the aDS for measurements of DA release with dLight1.3b. **b** Light (top) and tone (bottom) cue onset triggered averages of  $\Delta F/F$  for a single fiber (location in the coronal plane in inset) for trials where cues were associated with 100% (post learning phase), 90%, and 10% (extinction) probabilities. Shaded region, S.E.M. **c** Light cue-aligned  $\Delta F/F$  (top) and licking (bottom) for all trials for each phase included in the triggered averages in (b top). **d** Box plots showing the median DA  $\Delta F/F$  minima (dip, see Methods) for light (top) and tone (bottom) cue trials for trials post learning and the 10 and 90% reward probability phases during partial extinction. Each dot is a single fiber. \*\*\*:  $p < 0.001$  \*\*:  $p < 0.01$ , Kruskal-Wallis test. n.s.: not significant. **e** Top: mean light cue onset triggered DA ( $n = 29$  fibers across 4 mice) and ACh ( $n=68$  aDS fibers across 8 mice)  $\Delta F/F$  for cue onset during extinction (10% and 20% for DA and ACh, respectively). Shaded region, S.E.M. Bottom: Box plot of latencies to the minimum DA (DA dip) or maximum (ACh late peak) for all fibers in DA or ACh sensor-expressing mice, respectively, in the extinction phase. Each dot is one fiber. \* $p < 0.05$ , Two-tailed Wilcoxon rank-sum test. **f** Histogram showing the latencies across fibers of the emergence of significant DA dips to light (top) and tone (bottom) cues relative to the decrease in lick index following high to low reward probability transitions. Blue vertical lines indicate the medians of the distributions. **g** Box and whisker plots showing the peak decelerations and velocity changes across all mice ( $n=4$  mice) following light (left) and tone (right) cue onsets. **h** Left: Mean  $\Delta F/F$  for a fiber (location in inset) aligned to the onset of spontaneous, unrewarded spout licking bouts in the ITI on trials pre (blue) and post (orange) initial Pavlovian learning. Shaded region, S.E.M. Right: Lick bout-aligned  $\Delta F/F$  for all bout onsets included in triggered average on left. **i** Mean lick bout onset-triggered  $\Delta F/F$  across all mice and fibers ( $n = 29$  fibers across 4 mice) for bouts pre (blue) and post (orange) learning. Shaded region, S.E.M. **j** Box plots as in (d) showing the median DA  $\Delta F/F$  minima for spontaneous lick bout onsets pre- and post-learning. Each dot is one fiber. \*\* $p < 0.01$ , Two-tailed Wilcoxon rank-sum test. **k** Left: Mean lick bout onset-triggered DA ( $n = 29$  fibers across 4 mice) and ACh ( $n = 68$  aDS fibers across 8 mice)  $\Delta F/F$  in the aDS (see Methods) for bouts pre learning. Shaded region, S.E.M. Right: box plot of latencies to the minimum (DA dip) or maximum (ACh peak) for all fibers in DA or ACh sensor-expressing mice respectively in the pre learning phase. Each dot is one fiber. \*\*\* $p < 0.001$ , Two-tailed Wilcoxon rank-sum test.

Fig. 8



**1460 Fig. 8: Cue-evoked glutamate release onto anterior striatum cholinergic interneurons**  
**1461 does not change during extinction.** **a** Schematic showing strategy for single fiber optical  
1462 measurements of bulk glutamate release selectively on aDS CINs with iGluSnFR in ChAT-cre  
1463 mice. **b** Left: fluorescence image of a coronal section of the anterior striatum showing iGluSnFR  
1464 expression. Right: zoomed image of the region in the white box on left showing restricted  
1465 expression to cholinergic interneurons. **c** Representative  $\Delta F/F$  traces aligned to light cue onset  
1466 for two trials in one mouse. **d** Top: mean  $\Delta F/F$  from one representative fiber from a single mouse  
1467 aligned to light cue onset for trials post learning (orange) and partial extinction (20% reward  
1468 probability cue, green). Shaded regions, S.E.M. Bottom: light cue-aligned  $\Delta F/F$  for all trials  
1469 included in the triggered averages at top. **e** Mean  $\Delta F/F$  (left), licking (middle), and velocity (right)  
1470 ( $n = 8$  sessions across 2 fibers in 2 mice) aligned to light (top) and tone (bottom) cue onsets for  
1471 post learning and 20% partial extinction phases. Shaded regions, S.E.M.

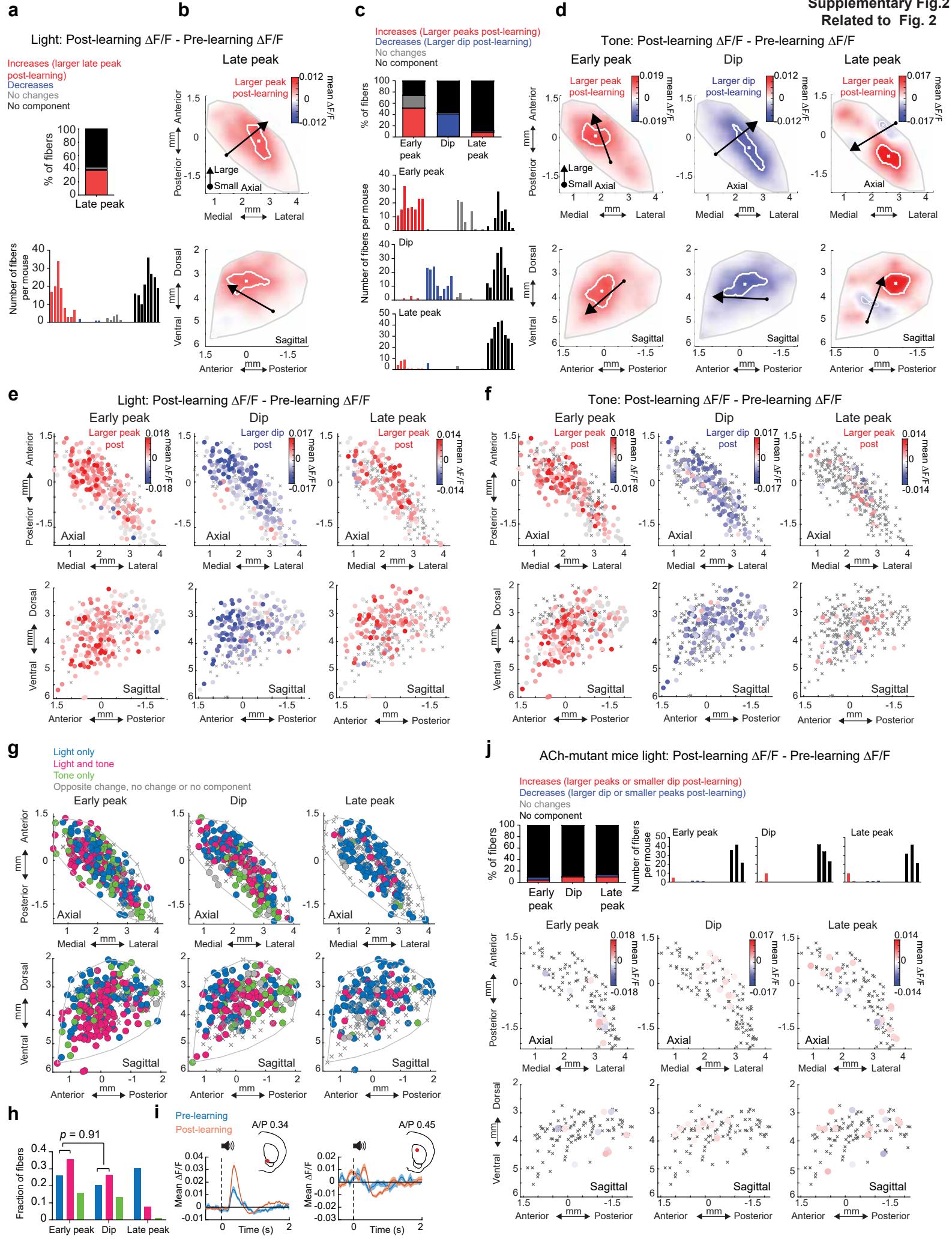
**Supplementary Fig. 1**  
Related to Fig. 2



## 1472 SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

1473 **Supplementary Fig. 1: Additional quantification of component specific changes in**  
1474 **cue-evoked ACh signals and mutant ACh-sensor controls during Pavlovian learning.** **a**  
1475 Histogram of latencies to significant (see Methods) peaks and troughs to light cue onset pre  
1476 (left) and post (right) learning. Shaded regions indicate time windows used to identify each  
1477 component (early peak, dip, and late peak), and the dashed lines indicate the means (pre  
1478 learning: early peak  $0.23 \pm 0.008$ s, dip  $0.47 \pm 0.009$ s, late peak  $0.82 \pm 0.045$ s; post learning: early  
1479 peak  $0.23 \pm 0.003$ s, dip  $0.52 \pm 0.005$ s, late peak  $0.93 \pm 0.012$ s) **b** Same as (**a**) for the tone cue (pre  
1480 learning: early peak  $0.32 \pm 0.01$ s, dip  $0.6 \pm 0.02$ s, late peak  $0.91 \pm 0.02$ s; post learning: early peak  
1481  $0.23 \pm 0.005$ s, dip  $0.52 \pm 0.009$ s, late peak  $0.93 \pm 0.03$ s). T-test for light vs. tone component  
1482 latencies: pre learning early peak:  $p = 9.3 \times 10^{-10}$ , post learning early peak  $p = 3.78 \times 10^{-19}$ ; pre  
1483 learning dip  $p = 2.07 \times 10^{-7}$ , post learning dip  $p = 1.97 \times 10^{-41}$ ; pre learning late peak  $p = 0.47$ , post  
1484 learning late peak  $p = 0.33$ . **c** Maps (axial left, sagittal right) indicating the presence of each  
1485 combination of signal components to light cue onset for each fiber (circle) pre (top) and post  
1486 (bottom) learning. Empty circles are fibers with no significant response. Pie charts indicate  
1487 percentages of fibers with each combination of signal components for light cue onset pre (top)  
1488 post (bottom) learning. **d** Same as (**c**) for the tone cue. **e** Mean light cue triggered  $\Delta F/F$  for post  
1489 learning trials for three representative fibers in the same mouse (locations in coronal sections at  
1490 bottom) illustrating the larger magnitude and longer latency ACh peaks in more ventral and  
1491 posterior locations and relatively larger dips in more anterior and dorsal regions, consistent with  
1492 the patterns shown in Fig. 2. **f** Maps (axial top, sagittal bottom) of latencies to early peaks to  
1493 light (left) and tone (right) cue onsets post learning for all fibers. Xs are fibers with no significant  
1494 early peak. **g** Light cue onset-triggered  $\Delta F/F$  averages (bottom) for 29 ventral striatum fibers in 8  
1495 mice on post-learning trials ( $n = 40-100$  trials). Colors of the traces correspond to the fiber  
1496 locations shown in the coronal plane schematics on top, where blue colors correspond to more  
1497 anterior locations and green colors to more posterior locations. Shaded regions, S.E.M.

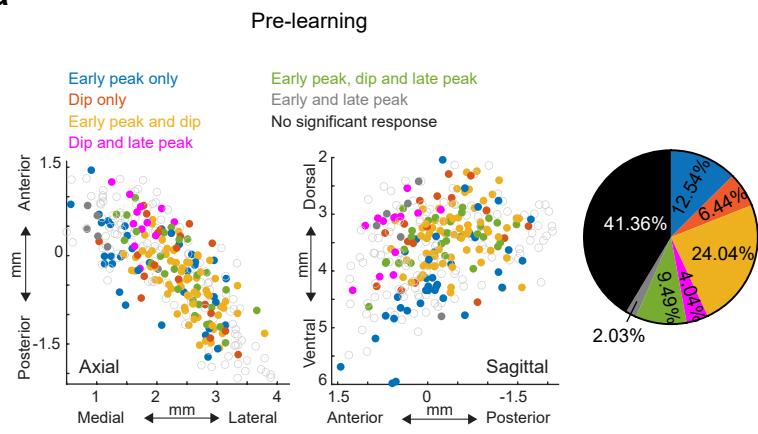
## Supplementary Fig.2 Related to Fig. 2



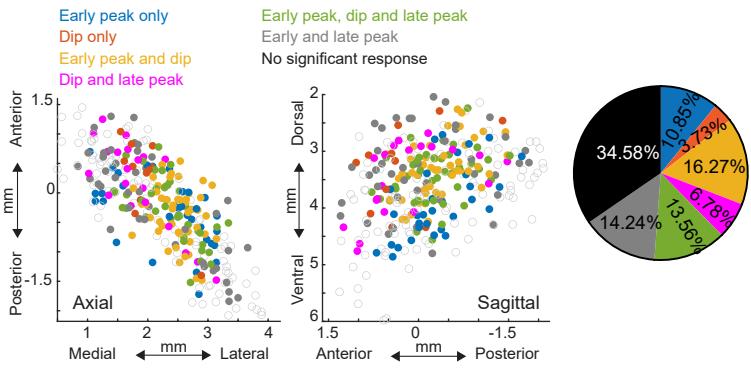
**1498 Supplementary Fig. 2: Additional quantification of the spatial organization of light and**  
**1499 tone cue-evoked ACh release changes during learning.** **a** Top: percent of all fibers with  
1500 significant increases or decreases, no change, or no significant component from pre to  
1501 post-learning for the late peak  $\Delta F/F$  at light cue onset. Bottom: histogram of number of fibers per  
1502 mouse with significant late peak changes. Each bar is the fiber count for one mouse for each  
1503 condition indicated by colors at left. **b** Maps (axial, top; sagittal, bottom) showing spatially  
1504 weighted means across locations of differences with learning (post - pre  $\Delta F/F$ ) for the mean late  
1505 peak  $\Delta F/F$  at light cue onset. Lines indicate the axes of maximal variation and arrows indicate  
1506 the direction of peak increases from smallest to largest changes. White contours indicate  
1507 regions with changes in the highest 10<sup>th</sup> percentile. **c** Top: percent of all fibers with significant  
1508  $\Delta F/F$  increases or decreases, no change, or no significant component from pre to post learning  
1509 for each component at tone cue onset. Bottom: histogram of number of fibers per mouse with  
1510 significant changes for each component. Each bar is the fiber count for one mouse for each  
1511 condition indicated by colors at top. **d** Maps (axial, top; sagittal, bottom) showing spatially  
1512 weighted means across locations of differences with learning (post-pre  $\Delta F/F$ ) for the mean  $\Delta F/F$   
1513 for the three signal components at tone cue onset. Lines indicate the axes of maximal variation  
1514 and arrows indicate the direction of components increases from smallest to largest changes.  
1515 White contours indicate regions with changes in the highest 10<sup>th</sup> percentile. **e** Maps showing the  
1516 difference in mean light-cue evoked  $\Delta F/F$  with learning (post – pre) for the three signal  
1517 components for each fiber (circle). Gray circles, no significant difference; Xs, no significant  
1518 component. Significance calculated with two-tailed Wilcoxon rank-sum test ( $p < 0.01$ ). **f** Same  
1519 as (e) for tone cue. **g** maps showing each fiber (dot) color coded according to whether  
1520 significant changes from pre to post learning were present for the light cue onset only (dark  
1521 blue), tone cue onset only (green) or both (pink). **h** Fraction of all fibers for each component  
1522 classified according to changes across learning for light cue, tone cue, or both as indicated in  
1523 (g) ( $p = 0.91$ , for early peak vs. dip, Fisher's exact test). **i** Mean  $\Delta F/F$  for 2 representative fibers  
1524 (same as the fibers shown in Fig. 2f (left) and Fig. 2h) aligned to tone cue onset for trials pre  
1525 (blue) and post (orange) learning. Shaded regions, S.E.M. Red dots in insets indicate the fiber  
1526 locations in the coronal plane. **j** Top left: percent of all fibers in mice expressing the  
1527 non-functional mutant ACh sensor with significant  $\Delta F/F$  increases or decreases, no change, or  
1528 no significant component from pre to post learning for each component at light cue onset. Top  
1529 right: histogram of number of fibers per mouse with significant early peak changes. Each bar is  
1530 the fiber count for one mouse for each condition indicated by colors at left. Bottom: Maps  
1531 showing the difference in mean light-cue evoked  $\Delta F/F$  with learning (post – pre) for the three  
1532 signal components for each fiber (circle) in mice expressing the mutant ACh sensor. Gray  
1533 circles, no significant difference; Xs, no significant component. Significance calculated with  
1534 two-tailed Wilcoxon rank-sum test ( $p < 0.01$ ).

**Supplementary Fig.3**  
Related to Fig. 3

**a**

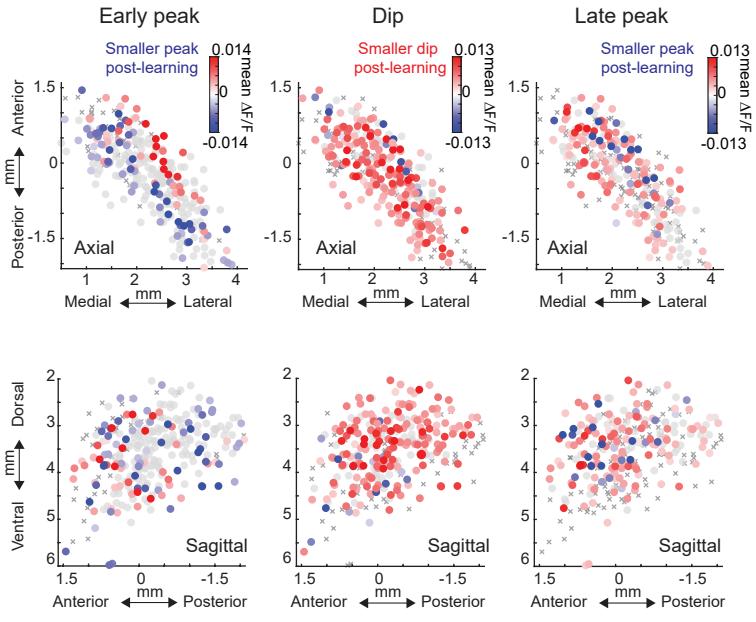


**Post-learning**



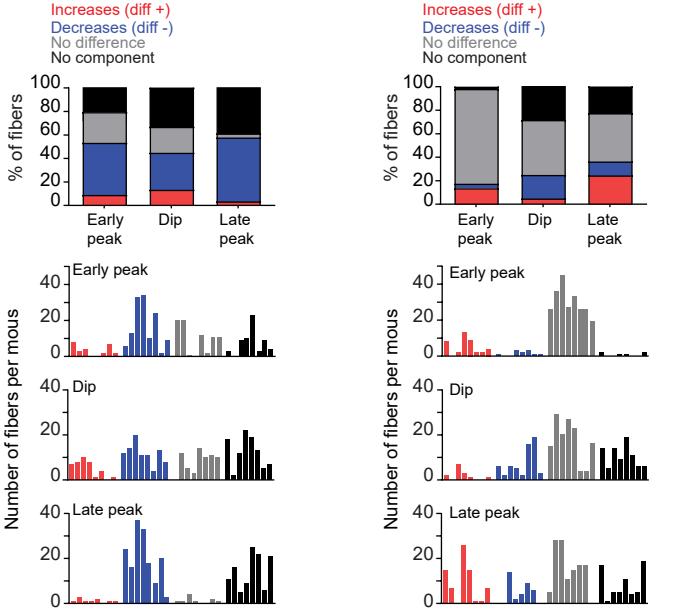
**b**

Reward: Post-learning  $\Delta F/F - \text{Pre-learning } \Delta F/F$



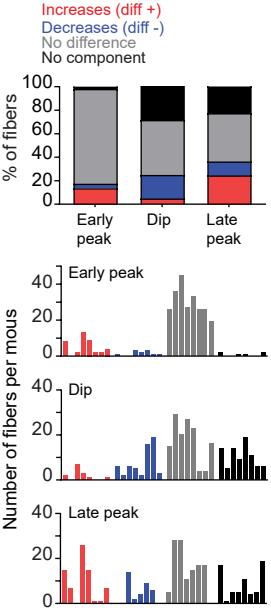
**c**

Unpredicted reward - post-learning reward



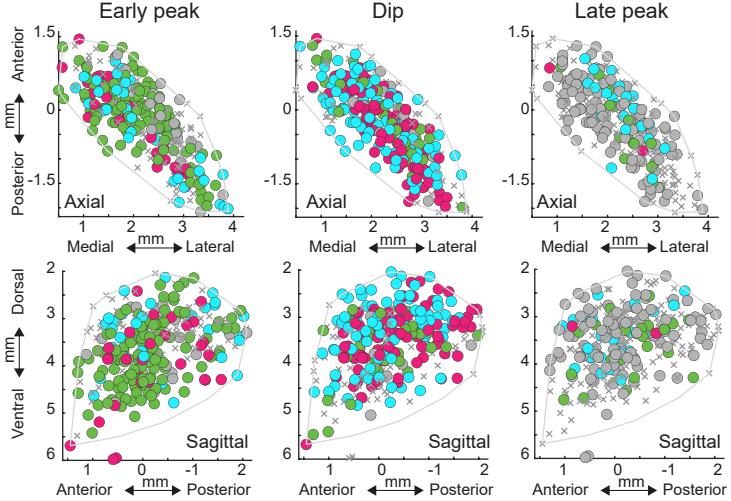
**d**

20% reward prob - 80% reward prob



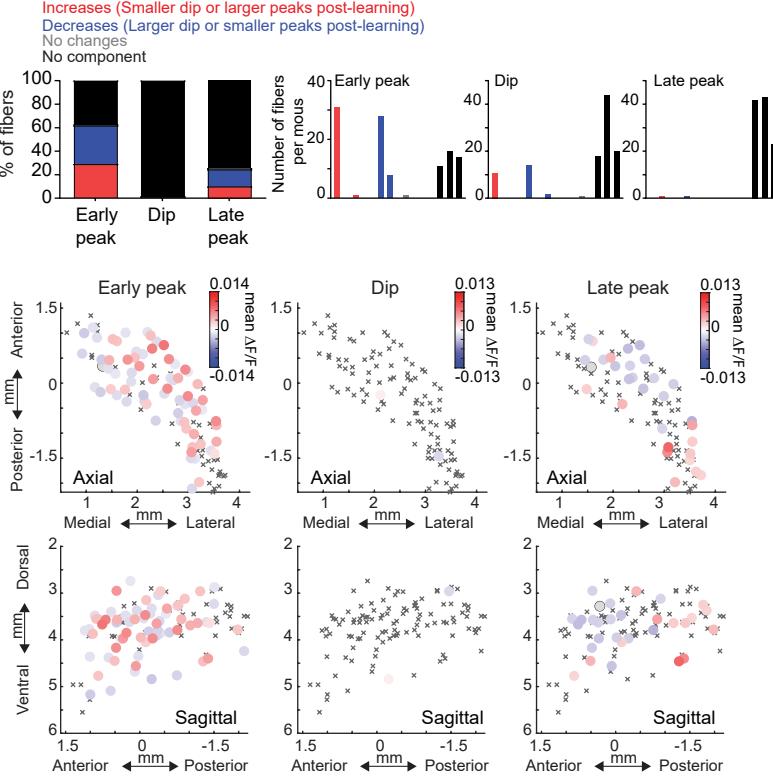
**e**

Tone only (green)  
Tone and reward (magenta)  
Reward only (cyan)  
Opposite change, no changes or no component (grey)



**f**

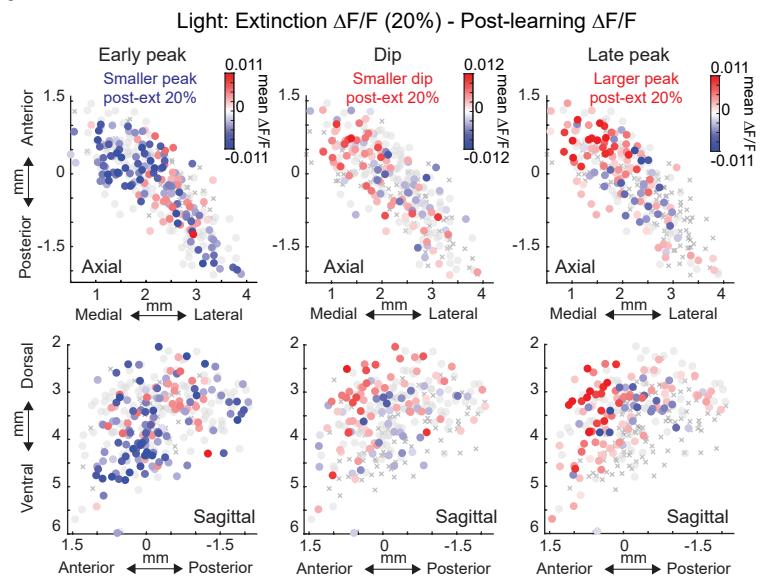
ACh mutant mice reward: Post-learning  $\Delta F/F - \text{Pre-learning } \Delta F/F$



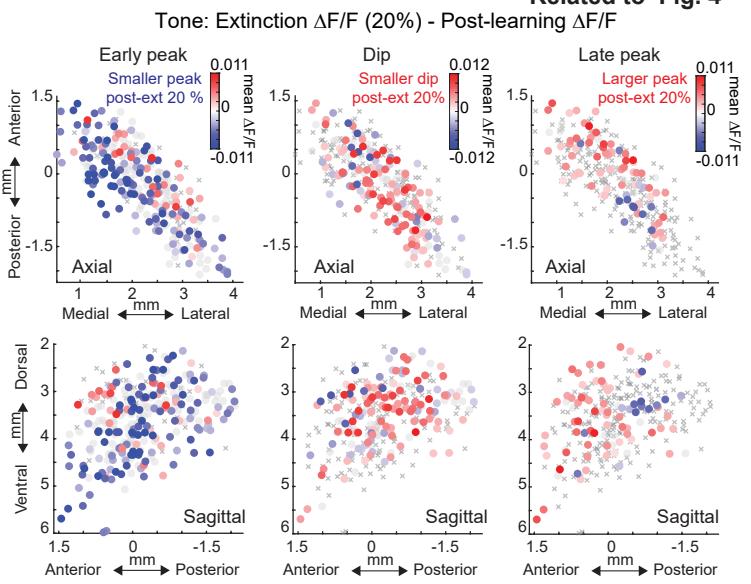
**1535 Supplementary Fig. 3: Additional quantification of changes in reward-evoked ACh release**  
**1536 and mutant ACh-sensor controls during learning.** **a** Maps indicating the presence of each  
1537 combination of signal components to reward consumption onset for each fiber (circle) pre (left)  
1538 and post (right) learning. Empty circles are fibers with no significant response. Pie charts  
1539 indicate percentages of fibers with each combination of significant signal components at reward  
1540 consumption onset pre (left) and post (right) learning. **b** Maps showing the difference in mean  
1541 reward consumption evoked  $\Delta F/F$  with learning (post – pre) for the three signal components for  
1542 each fiber (circle). Gray circles, no significant difference; Xs, no significant component.  
1543 Significance calculated with two-tailed Wilcoxon rank-sum test ( $p < 0.01$ ). **c** Top: percent of all  
1544 fibers with significant  $\Delta F/F$  differences between unpredicted random rewards and cue predicted  
1545 rewards post-learning for each component. Bottom: histogram of number of fibers per mouse  
1546 with significant changes for each component. Each bar is the fiber count for one mouse for each  
1547 condition indicated by colors at top. **d** Same as (c) for  $\Delta F/F$  differences between rewards  
1548 following the 20% (less predicted) and 80% (more predicted) cues during the extinction phase. **e**  
1549 Maps showing each fiber (circle) color coded according to whether significant changes from pre  
1550 to post learning were present at the tone cue onset only (green), reward consumption only (light  
1551 blue) or both (pink). Pink dots indicate locations where the component magnitude became larger  
1552 (dip more negative or peak more positive) for cue and smaller for reward over learning,  
1553 consistent with reward prediction error encoding. **f** Maps as in (b) showing the difference in  
1554 mean reward evoked  $\Delta F/F$  with learning (post – pre) for the three signal components for each  
1555 fiber (circle) in mice expressing the mutant ACh sensor. Gray circles, no significant difference;  
1556 Xs, no significant component.. Significance calculated with two-tailed Wilcoxon rank-sum test ( $p$   
1557  $< 0.01$ ). **g** Left: percent of all fibers in mice expressing the non-functional mutant ACh sensor  
1558 with significant  $\Delta F/F$  increases or decreases, no change, or no significant component from pre  
1559 to post-learning for each component at reward consumption onset. Right: histogram of number  
1560 of fibers per mouse with significant changes. Each bar is the fiber count for one mouse for each  
1561 condition indicated by colors at left. Note that a larger fraction of artifacts were present for the  
1562 early peak, but these changes were relatively small, were predominantly present in one mouse,  
1563 and did not correspond with the primary changes highlighted in the ACh3.0 sensor mice (see  
1564 Fig. 3).

**Supplementary Fig. 4**  
Related to Fig. 4

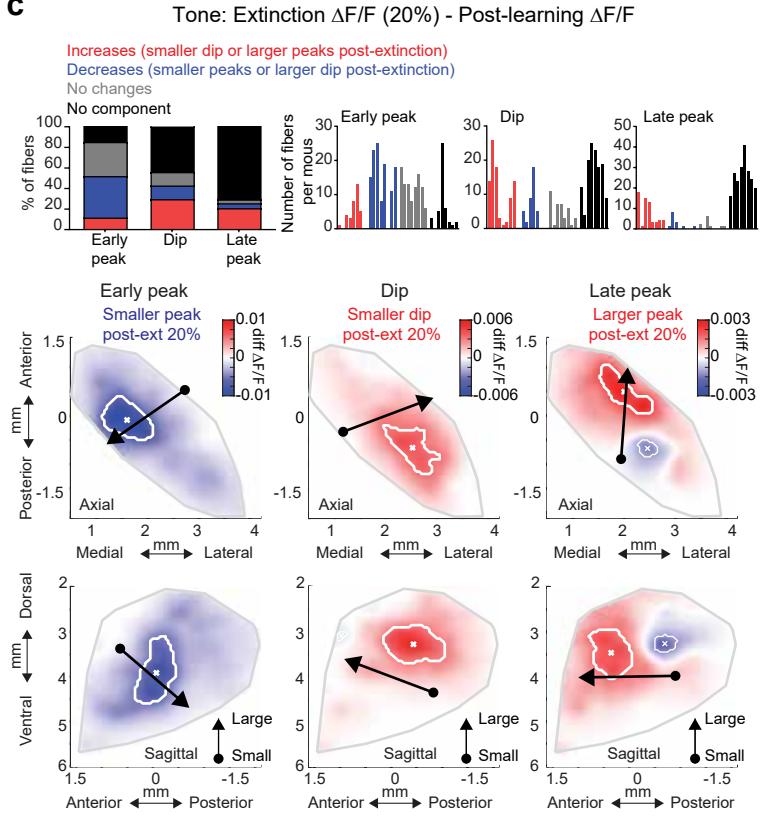
**a**



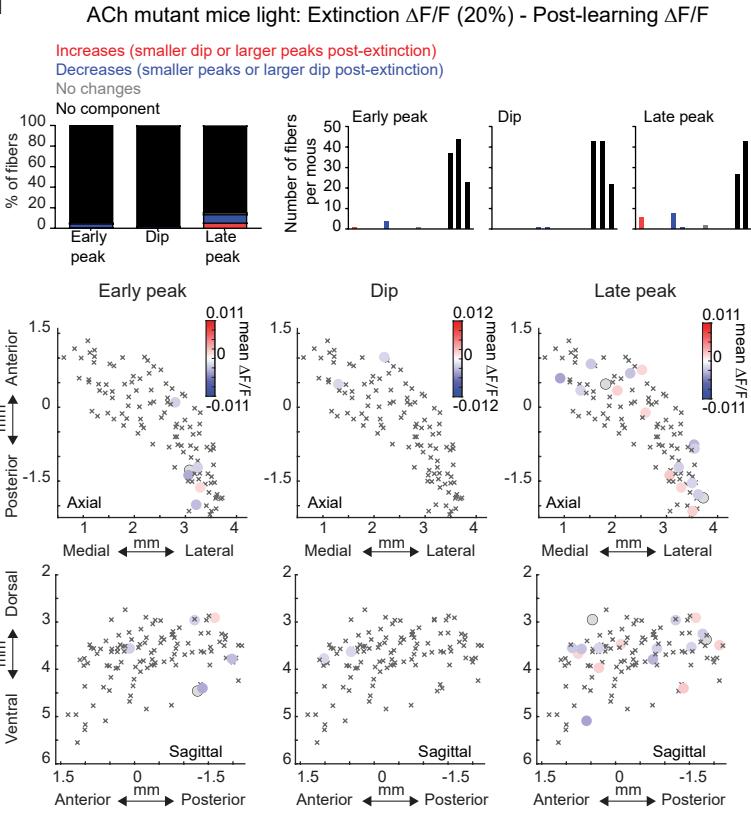
**b**



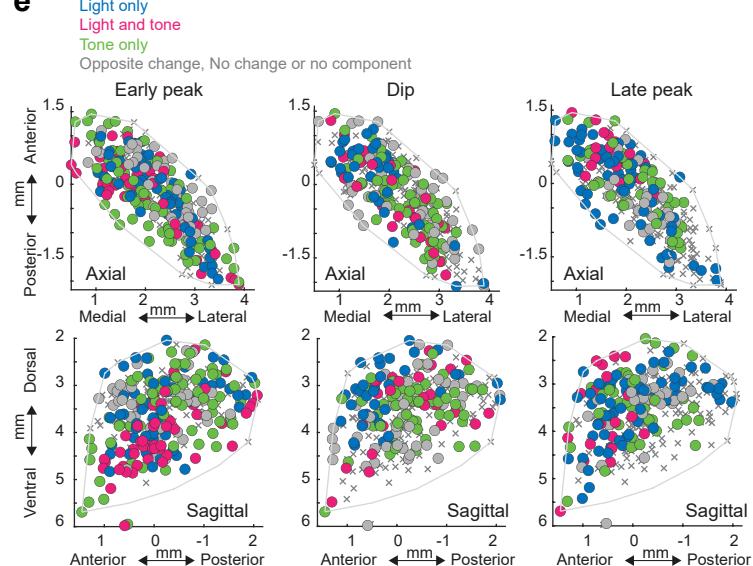
**c**



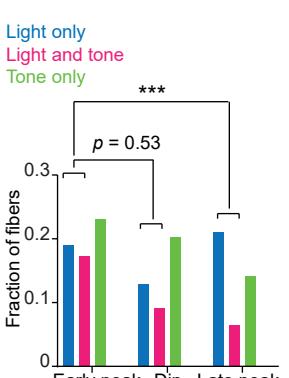
**d**



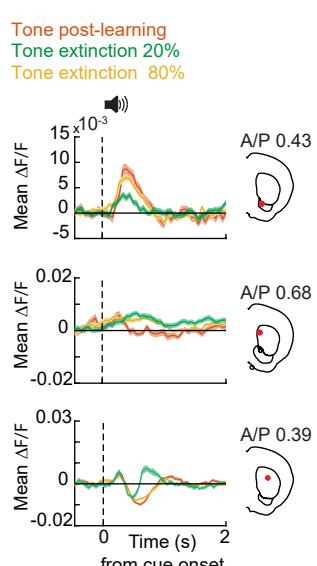
**e**



**f**



**g**

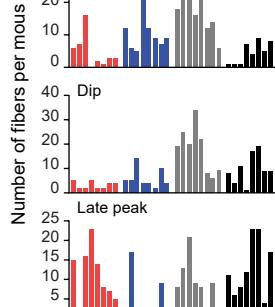
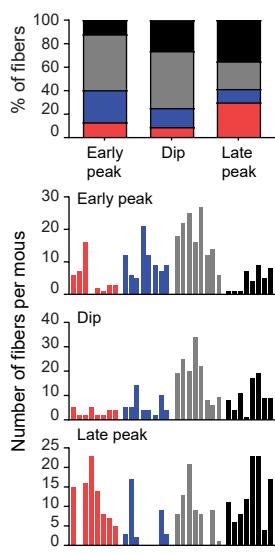


**1565 Supplementary Fig. 4: Additional quantification of changes in cue-evoked ACh release**  
**1566 and mutant ACh-sensor controls during extinction.** **a** Maps showing the difference in mean  
1567 light cue-evoked  $\Delta F/F$  between the post-learning (100%) and 20% reward probability phases for  
1568 the three signal components. Each circle is one fiber. Gray circles, no significant difference; Xs,  
1569 no significant component. Significance calculated with two-tailed Wilcoxon rank-sum test ( $p <$   
1570 0.01). **b** Same as **(a)** for tone. **c** Top left: percent of all fibers with significant  $\Delta F/F$  differences  
1571 between the 20% extinction and 100% post learning phases (20% - 100%  $\Delta F/F$ ) at tone cue  
1572 onset for each component. Top right: histogram of number of fibers per mouse with significant  
1573 changes for each component. Each bar is the fiber count for one mouse for each condition  
1574 indicated by colors at top. Bottom: Maps (axial, top; sagittal, bottom) showing spatially weighted  
1575 means across locations of differences between the 20% extinction and 100% post learning  
1576 phases (20% - 100%  $\Delta F/F$ ) for the mean early peak  $\Delta F/F$  at tone cue onset. Lines indicate the  
1577 axes of maximal variation and arrows indicate the direction of peak decreases from smallest to  
1578 largest changes. White contours indicate regions with changes in the highest 10th percentile. **d**  
1579 Top: fiber percentages and counts per mouse as in **(c)** for the light cue in mice expressing the  
1580 non-functional mutant ACh sensor. Bottom: maps as in **(a)** for the mutant ACh sensor mice for  
1581 the light cue. **e** Maps showing each fiber (circle) color coded according to whether significant  
1582 differences between the 20% extinction and 100% post learning phases were present for the  
1583 light cue onset only (dark blue), tone cue onset only (green) or both (pink). **f** Fraction of all fibers  
1584 for each component classified according to changes across extinction for light cue, tone cue, or  
1585 both as indicated at top ( $p = 0.53$  for early peak vs. dip and  $p = 0.0008$  for early peak vs. late  
1586 peak; Fisher's exact test, \*\*\*  $p < 0.001$ ). **g** Mean  $\Delta F/F$  for 3 representative fibers aligned to tone  
1587 cue onset for trials where tone was associated with 100% (post-learning), 80%, and 20%  
1588 (extinction) probabilities. Shaded regions, S.E.M. across trials. Red dots in the insets indicate  
1589 the fiber locations in the coronal plane. Top and middle examples are the same fibers shown in  
1590 Fig. 4e and g, respectively).

**Supplementary Fig.5**  
Related to Fig. 4

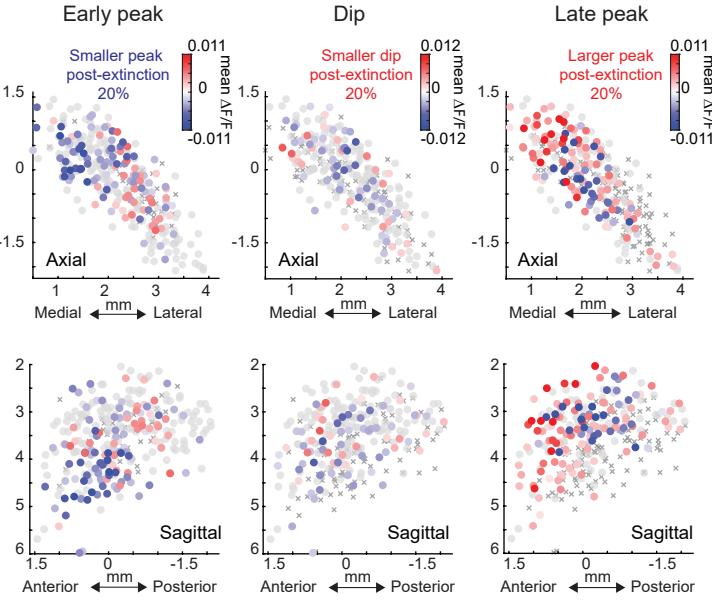
**a**

Increases (smaller dip or larger peaks post-extinction 20%)  
Decreases (smaller peaks or larger dip post-extinction 20%)  
No changes  
No component



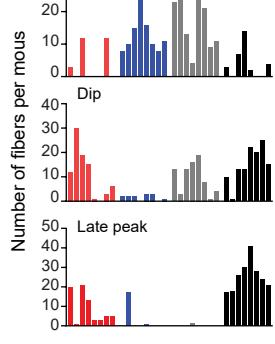
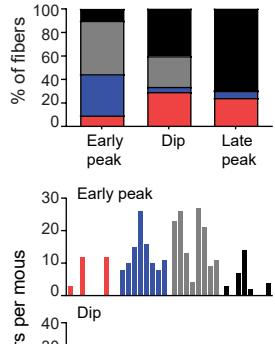
**b**

Light: Extinction  $\Delta F/F$  (20%) - Extinction  $\Delta F/F$  (80%)



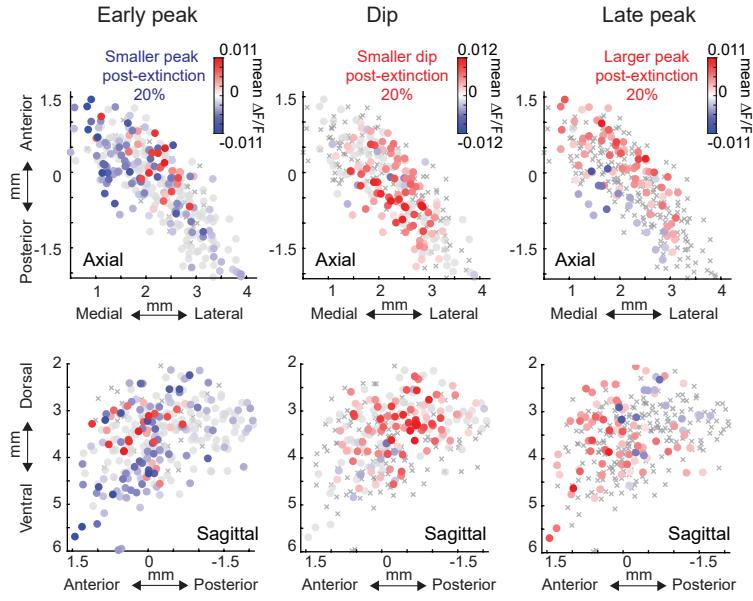
**c**

Increases (smaller dip or larger peaks post-extinction 20%)  
Decreases (smaller peaks or larger dip post-extinction 20%)  
No changes  
No component



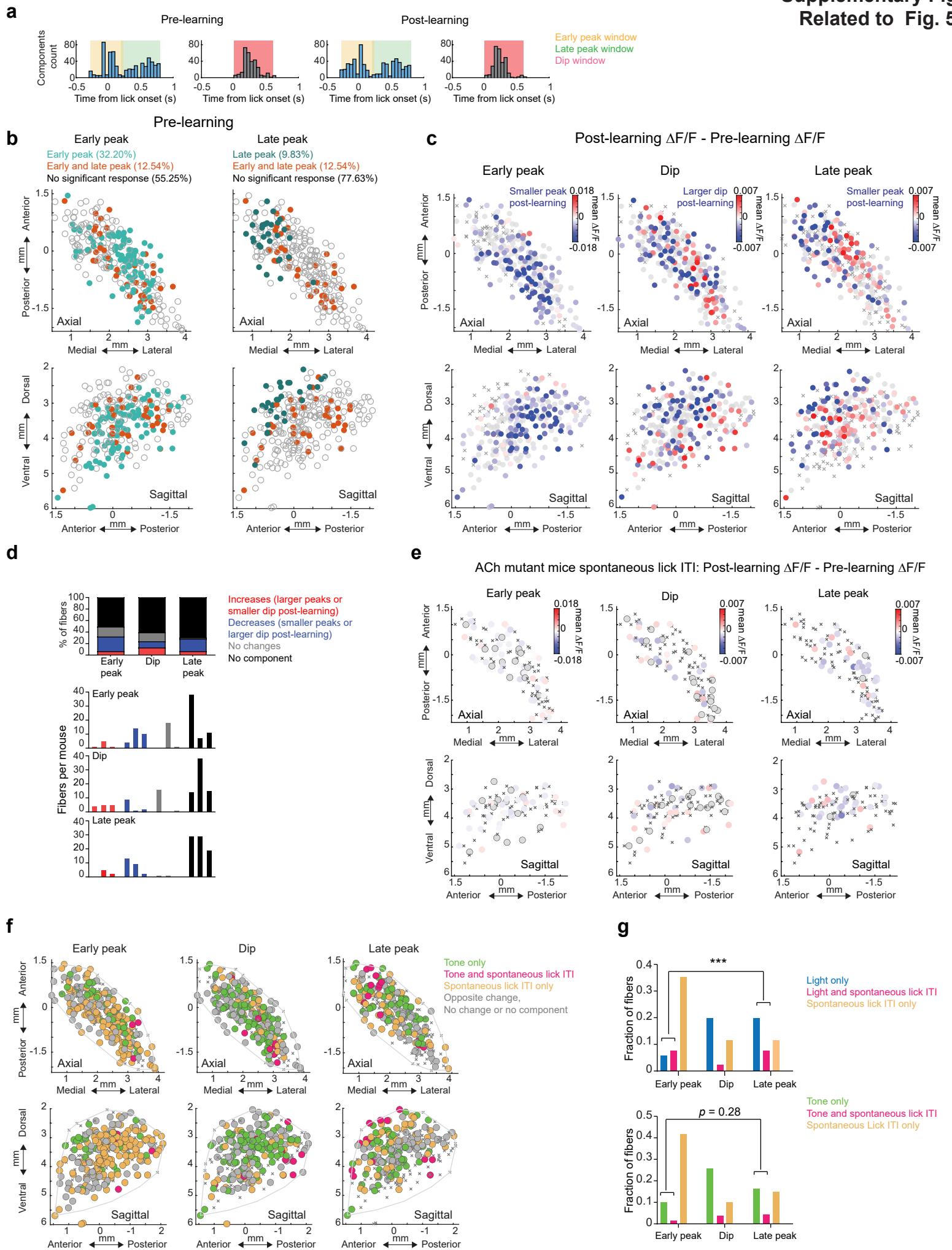
**d**

Tone: Extinction  $\Delta F/F$  (20%) - Extinction  $\Delta F/F$  (80%)



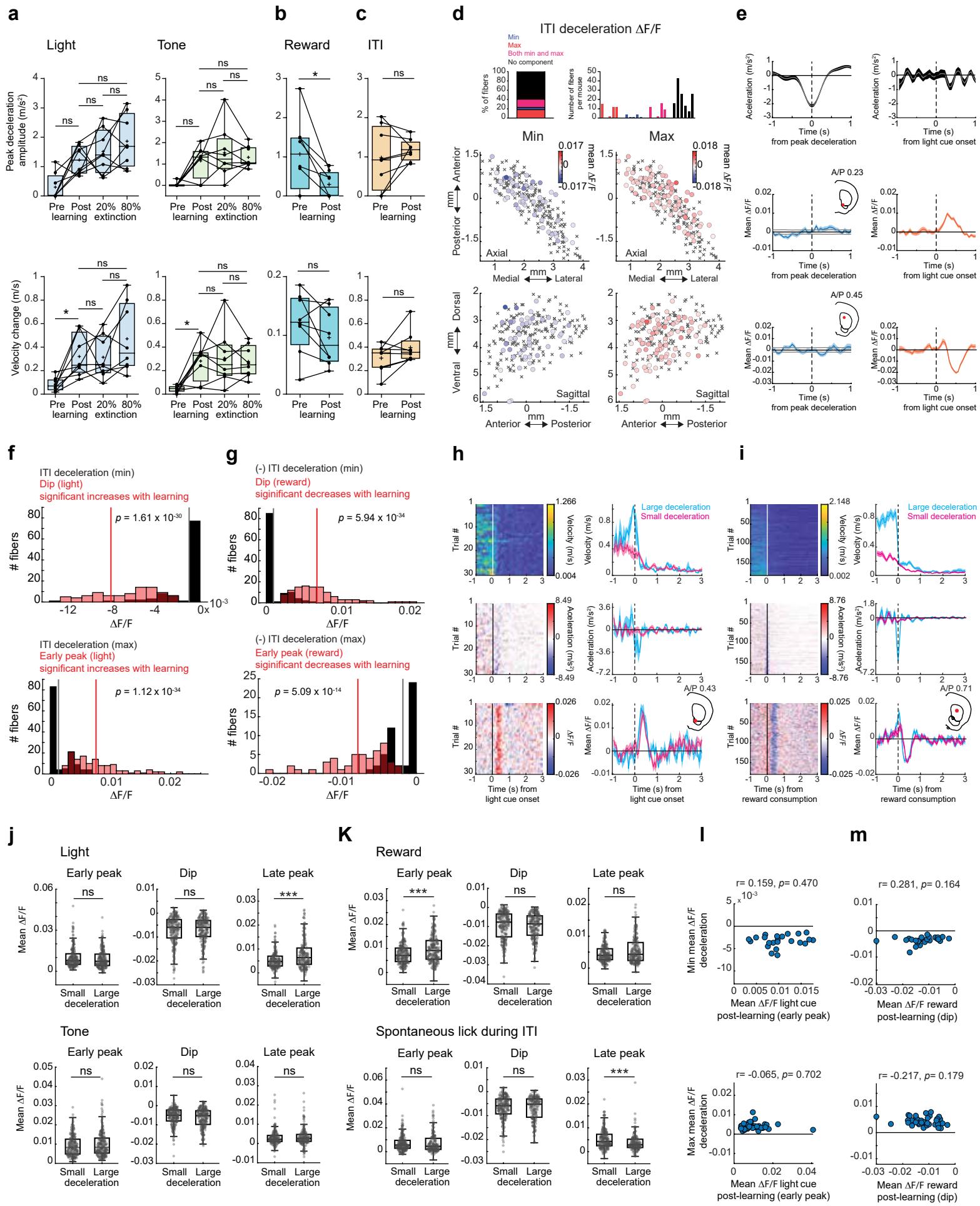
**1591 Supplementary Fig. 5: Comparison of ACh release for high and low probability cues**  
**1592 during extinction sessions.** **a** Top: percent of all fibers with significant  $\Delta F/F$  differences  
1593 between the 20% extinction and 80% extinction phases (20% - 80%  $\Delta F/F$ ) at light cue onset for  
1594 each component. Bottom: histogram of number of fibers per mouse with significant changes for  
1595 each component. Each bar is the fiber count for one mouse for each condition indicated by  
1596 colors at top. **b** Maps showing the difference in mean light cue-evoked  $\Delta F/F$  between the 80%  
1597 and 20% reward probability phases during extinction (20% - 80%) for the three signal  
1598 components. Each circle is one fiber. Gray circles, no significant difference; Xs, no significant  
1599 component. Significance calculated with two-tailed Wilcoxon rank-sum test ( $p < 0.01$ ). **c**  
1600 Percentages and fiber counts per mouse as in (a) for  $\Delta F/F$  differences between the 20%  
1601 extinction and 80% extinction phases (20% - 80%  $\Delta F/F$ ) at tone cue onset. **d** Maps as in (b) for  
1602 the tone cue.

**Supplementary Fig. 6**  
Related to Fig. 5



**1603 Supplementary Fig. 6: Additional quantification of changes in unrewarded lick evoked**  
**1604 ACh release and mutant ACh-sensor controls during learning.** **a** Histogram of latencies to  
1605 significant (see Methods) peaks and troughs to unrewarded ITI lick bout onsets pre and post  
1606 learning. Shaded regions indicate time windows used to identify each component (early peak,  
1607 dip, and late peak). **b** Maps (axial, left; sagittal, right) indicating the presence of each  
1608 combination of signal components to unrewarded ITI lick bouts for each fiber (circle) in the  
1609 pre-learning phase. Empty circles are fibers with no significant response. **c** Maps showing the  
1610 difference in mean ITI lick-evoked  $\Delta F/F$  between the pre and post learning phases (pre - post)  
1611 for the three signal components. Each circle is one fiber. Gray circles, no significant difference;  
1612 Xs, no significant component. Significance calculated with two-tailed Wilcoxon rank-sum test ( $p$   
1613  $< 0.01$ ). **d** Top: percent of all fibers with significant  $\Delta F/F$  differences between the pre and post  
1614 learning phases at ITI lick onset for each component in mice expressing non-functional mutant  
1615 ACh sensor. Bottom: histogram of number of fibers per mouse with significant changes for each  
1616 component. Each bar is the fiber count for one mouse for each condition indicated by colors at  
1617 top. **e** Same as (b) for mice expressing the non-functional mutant ACh sensor. Note that  
1618 although some sparse artifacts were detected in mutant sensor mice, the spatial organization  
1619 and magnitude of the artifacts did not match the true signals in the ACh sensor mice (Fig. 5). **f**  
1620 Maps (axial, top; sagittal, bottom) showing each fiber (dot) color coded according to whether  
1621 significant changes from pre to post learning (see Methods) were present at the tone cue onset  
1622 only (dark blue), unrewarded ITI lick only (light blue) or both (pink). Pink dots indicate locations  
1623 where the component magnitude was larger with extinction (dip more negative or peak more  
1624 positive) for cue and smaller post-learning for the unrewarded ITI lick, consistent with negative  
1625 reward prediction error encoding. **g**, Fraction of all fibers for each component classified  
1626 according to changes across extinction for light cue (top) or tone cue (bottom) and across  
1627 learning for unrewarded licks as indicated in (f).  $p = 0.0027$  for early peak vs. late peak,  $p = 0.28$   
1628 for early peak vs. late peak, \*\*\* $p < 0.001$ , Fisher's exact test).

**Supplementary Fig. 7**  
Related to Fig. 2,3,4 and 5



**1629 Supplementary Fig. 7: Learning and extinction related changes in striatum-wide ACh**  
**1630 release cannot be explained by locomotion changes.** **a** Box and whisker plots showing  
1631 average peak decelerations (top) and velocity changes (bottom) across all mice (n=8 mice)  
1632 following light (left) tone (right) cue onsets for each learning phase. (n.s., not significant  
1633 Friedman test,  $p > 0.05$ ,  $p = 0.04$  for light and tone velocity,  $*p < 0.05$ ). Each datapoint is the  
1634 mean for one mouse. **b** Same as **(a)** for reward delivery pre and post learning (Two-tailed  
1635 Wilcoxon rank-sum test,  $p = 0.03$  for peak deceleration and  $p = 0.14$  for velocity;  $*p < 0.05$ ). **c**  
1636 Same as **(b)** for unrewarded lick bout onsets in the ITI (Two-tailed Wilcoxon test,  $p = 0.38$  for  
1637 peak deceleration and  $p = 0.19$  for velocity). **d** Top left: Percent of all fibers with significant  
1638 minima, maxima, or both (min and max) of the mean ACh  $\Delta F/F$  within a  $\pm 1$ -s window triggered  
1639 on troughs of large spontaneous decelerations in the intertrial interval (ITI). Top right: Histogram  
1640 of the number of fibers per mouse with significant changes for minima, maxima, or both. Bottom:  
1641 Maps (axial, top; sagittal, bottom) showing the minima (left) and maxima (right) of the mean ACh  
1642  $\Delta F/F$  triggered on troughs of large ITI decelerations. Significance for each fiber (circle) was  
1643 determined by comparison to the 95% confidence interval of a bootstrapped  $\Delta F/F$  distribution  
1644 (Methods). Xs, no significant response. **e** Left column: Mean acceleration (top) and mean  $\Delta F/F$   
1645 (middle and bottom) aligned to ITI deceleration troughs for 2 representative fibers with small but  
1646 significant minima and maxima. Shaded regions, S.E.M. Gray shading indicates the 95 %  
1647 confidence interval. Insets indicate fiber location in the coronal plane. Right column: Mean  
1648 acceleration (top) and mean  $\Delta F/F$  (middle and bottom) aligned to light cue onset for trials  
1649 post-learning for the same representative fibers shown in the left column. Shaded region,  
1650 S.E.M. **f** Red: Histogram showing the magnitudes of  $\Delta F/F$  changes from pre to post learning for  
1651 the dip (top) and early peak (bottom) components for the light cue for fibers with significant  
1652 changes. Black: Histogram of the magnitudes of the  $\Delta F/F$  minima (top) and maxima (bottom) of  
1653 the spontaneous ITI deceleration triggered  $\Delta F/F$  for the same fibers with changes over learning  
1654 plotted in red. Fibers with no significant deceleration ITI response counted in the 0 bin. Lines  
1655 indicate means (Two-tailed Wilcoxon rank-sum test). **g** Same as **(f)** for reward delivery periods.  
1656 The sign of the ITI deceleration response magnitudes is flipped for visualization because the  
1657 deceleration magnitudes to reward delivery decrease with learning, opposite to cues. **h** Left  
1658 column: Total treadmill velocity (top), acceleration (middle), and  $\Delta F/F$  (bottom) aligned to light  
1659 cue onsets for all trials post-learning for a single example fiber with a significant early peak  
1660 increase with learning sorted by the deceleration magnitude following cue onset. Right column:  
1661 Mean cue-triggered averages for measures in left column for trials with large decelerations  
1662 ( $>80$ th percentile of decelerations from the distribution of decelerations across all mice, n = 4  
1663 trials, blue) and small decelerations ( $<20$ th percentile, n = 8 trials, magenta). Shaded region,  
1664 S.E.M. **i** Same as **(j)** but aligned on reward consumption for another example fiber with a  
1665 significant change in the reward dip magnitude with learning. (n = 24 and 31 trials for large and  
1666 small deceleration trials respectively). **j** Box and whisker plots showing the mean light (top) and  
1667 tone (bottom) cue-evoked  $\Delta F/F$  for each component (early peak, dip and late peak) for trials  
1668 with small and large decelerations across all fibers (dots). Box bounds indicate the 25<sup>th</sup> and 75<sup>th</sup>  
1669 percentiles and the whiskers represent the minima and maxima. Significance calculated  
1670 between large and small decelerations for each component across fibers using a two tailed  
1671 Wilcoxon rank-sum test. \*\*\*:  $p < 0.001$ . n.s.: not significant. **k** same as **(j)** for reward (top) and  
1672 spontaneous ITI licks (bottom). **l** Scatterplots of the minimum (top) and maximum (bottom) of  
1673 the mean  $\Delta F/F$  triggered on ITI deceleration troughs and the mean post-learning light  
1674 cue-evoked early peak  $\Delta F/F$  for all fibers (dots) with significant components for both cues and  
1675 ITI decelerations. R and p values are from Pearson's correlations. **m** same as **(l)** but for the dip  
1676 component during reward post-learning.

Table 1: Pearson's correlation of minimum and maximum mean  $\Delta F/F$  of ITI deceleration with cue and reward component signals

		Early Peak		Dip		Late peak	
		r	p-value	r	p-value	r	p-value
Light pre-learning	Min						
	Max						
Light post-learning	Min	0.159	0.47	0.497	0.042	0.315	0.346
	Max	-0.065	0.702	-0.159	0.429	0.307	0.201
Tone pre-learning	Min						
	Max	0.398	0.178	0.868	0.056		
Tone post-learning	Min	-0.324	0.093	0.074	0.802		
	Max	0.245	0.097	-0.565	0.006		
Reward pre-learning	Min	0.022	0.923	-0.342	0.094	-0.274	0.242
	Max	-0.022	0.902	-0.325	0.038	0.088	0.668
Reward post-learning	Min	0.292	0.24	0.281	0.164	0.332	0.113
	Max	0.092	0.61	-0.217	0.179	-0.058	0.737

**1677 Table 1:** Pearson's correlation coefficients and *p*-values between the minimum and maximum of  
1678 the mean  $\Delta F/F$  triggered on ITI deceleration troughs and the mean cue or reward evoked  $\Delta F/F$   
1679 for the three components for pre and post learning phases. Correlations are limited to fibers with  
1680 both significant deceleration and task-related changes for each component. Empty cells are  
1681 conditions with fewer than 10 fibers.