Spatially organized striatum-wide acetylcholine dynamics for the learning and extinction of Pavlovian cues and actions

Safa Bouabid[1], Liangzhu Zhang[1], Mai-Anh T. Vu[1], Kylie Tang[1], Benjamin M. Graham[1], Christian A. Noggle[1], Mark W. Howe[1,*]

[1]Department of Psychological & Brain Sciences, Boston University, Boston, MA, USA
[*]Corresponding author: (mwhowe@bu.edu)

Abstract

Striatal acetylcholine (ACh) has been linked to behavioral flexibility. A key component of flexibility is down-regulating responding as valued cues and actions become decoupled from positive outcomes. We used array fiber photometry in mice to investigate how ACh release across the striatum evolves during learning and extinction of Pavlovian associations. Changes in multi-phasic release to cues and consummatory actions were bi-directional and region-specific. Following extinction, increases in cue-evoked ACh release emerged in the anterior dorsal striatum (aDS) which preceded a down-regulation of anticipatory behavior. Silencing ACh release from cholinergic interneurons in the aDS blocked behavioral extinction. Dopamine release dipped below baseline for down-shifted cues, but glutamate input onto cholinergic interneurons did not change, suggesting an intrastriatal mechanism for the emergence of ACh increases. Our large-scale mapping of striatal ACh dynamics during learning pinpoints region-specific elevations in ACh release positioned to down-regulate behavior during extinction, a central feature of flexible behavior.

Introduction

Cholinergic interneurons (CINs) provide the main source of acetylcholine (ACh) to the striatum[1,2]. Selective manipulations in rodents have revealed potential roles for the CINs in different forms of behavioral flexibility, response suppression, and extinction, which vary by striatal subregion[3–13]. Measurements of CIN activity or ACh release in the striatum during behavior have focused extensively on dynamics to salient and reward associated sensory stimuli[14,15]. Early electrophysiological recordings of putative CINs (tonically active neurons) in the primate caudoputamen revealed a characteristic multi-phasic response profile to cues and rewards consisting of a short latency burst in firing, followed by a pause, then a post-inhibitory rebound[16–18]. Each of these components, particularly the pause, are variably expressed and can evolve during associative learning. Pause amplitude and duration increase as cues become associated with reward, and these changes can be modulated by coincident elevations in dopamine (DA) during learning[17,19,20]. Similar multi-phasic dynamics and learning related changes are present in rodent CINs and striatal ACh release[21–24]. Mechanistic studies in brain slices and anesthetized rodents have indicated that pauses in ACh release to conditioned cues which coincide with phasic DA release may promote associative learning by creating a permissive window for striatal synaptic plasticity on D1 receptor expressing striatal projection neurons[25–27]. Less is known about how endogenous ACh dynamics to cues or actions might support a role in behavioral flexibility, but

there is some evidence that increases in ACh may weaken excitatory striatal synapses to facilitate extinction[9,13].

Functional heterogeneities in ACh signaling across striatal subregions have been difficult to assess due to the sparseness of CINs in the striatum (~3-5% of total population) and inherent limitations in spatial coverage of current optical and electrophysiological approaches. Thus, little is known about how ACh signaling varies across striatal subregions to support distinct roles in learning and flexibility. CIN recordings in primates report relatively uniform responses across putative CINs[15,17,19], while recent studies in rodents indicate that there may be differences in the expression of the tri-phasic signals across the dorsal/ventral striatal axis[22,23,28–31]. To overcome these limitations, we applied a multi-fiber array technique to measure rapid ACh release dynamics across >50 locations simultaneously throughout the entire 3-dimensional striatum volume in head-fixed, behaving mice. These measurements have revealed novel, region-specific, event-locked increases in ACh that are positioned to enable flexible shifts in behavior by modulating the extinction of actions and learned cue values.

Results

We optically measured rapid ACh release dynamics across the striatum in head-fixed behaving mice using large-scale arrays of 55-99 optical micro-fibers (50µm diameter) distributed uniformly across the 3-dimensional striatum volume (**Fig. 1A**, see Methods). Each fiber collected fluorescence from post-synaptic neurons expressing the genetically encoded ACh sensor, GRAB-ACh3.0[32] within a tapered collection volume extending approximately 100µm axially and 25µm radially from each fiber tip (**Fig. 1A-C**,[33]). Fiber spacing in all dimensions ensured no overlap in the collection volumes for each fiber (see Methods, [33–35]). Fiber locations in the brain were precisely reconstructed from post-mortem computerized tomography scans and aligned to a common coordinate framework to enable comparisons of spatial patterns of ACh release within and across subjects (**Fig. 1D-E**, [36,37], see Methods). Rapid, bi-directional changes in fluorescence, a proxy for changes in ACh release and synchronous firing of cholinergic interneurons[38], could be reliably detected across striatum locations (**Fig. 1C**). To control for potential signal contamination by hemodynamic or motion artifacts, we conducted measurements with quasi-simultaneous 405 illumination, the approximate isosbestic point of the sensor[32] and in mice expressing a non-functional null version of the ACh3.0 sensor (**Fig. 1B**, **Supplementary Fig.1**). Rapid peaks and dips were generally not observed in these control conditions. However, some slow, low amplitude fluctuations were present, which could largely be eliminated with high pass filtering (see Methods, **Supplementary Fig. 1D**). In mice expressing the null sensor, we found that some small artifacts were present with 470nm illumination which were not detected with 405nm illumination (**Supplementary Fig. 1B-C**). Therefore, quasi-simultaneous 405nm illumination may not reliably account for all potential signal artifacts, likely because of differences in hemoglobin absorbance and tissue scattering between the two illumination wavelengths[39]. Thus, going forward, we used the null sensor expressing cohort with 470nm illumination to rule out significant contributions of artifacts to behavior related signals in ACh3.0 expressing mice.

Unpredicted water reward deliveries produced multi-phasic release profiles at some locations, consisting of an early peak, followed by a dip below baseline, then a later peak (**Fig. 1F-G**).

However, all three signal components were not present in every region (**Fig. 1H**). Dips were present across nearly the entire dorsal striatum but were absent in the posterior ventral striatum, where only the early peak component was present. Early peaks were also observed at many, but not all locations across the dorsal striatum. The magnitudes of each component were generally stable across days enabling us to account for potential changes in sensor expression or light collection during chronic recordings. Only a small fraction of fibers showed a sparse statistically significant changes in unpredicted reward mean $\Delta F/F$ in ACh3.0 expressing mice across days of learning (5, 17 and 17 fibers out of a total of 295 fibers within the striatum for early peak, dip and late peak, respectively, $p<0.01$, one-way ANOVA test comparing the mean of each day with the mean of every other day (see Methods, **Fig. 1I**). Thus, our novel optical approach enabled us to investigate the evolution of simultaneous, spatiotemporal patterns of ACh release across the striatum during learning.

Region specific changes in striatal ACh release during Pavlovian learning

To determine how and where learning related changes in cue evoked ACh release dynamics occur, we mapped changes in ACh release across the striatum as mice learned a dual-cue delay Pavlovian conditioning task. Head-fixed mice (n = 8) on a spherical treadmill[40] were trained to associate light (470nm LED, 7mW) and tone (12kHz, 80dB) stimuli with water reward after a fixed 3s delay (**Fig. 2A**). Cues were presented one at a time and pseudo randomly interleaved with 4-40s intertrial intervals (ITIs) drawn from a uniform distribution. Non-contingent rewards were delivered occasionally (8 per session) in ITI periods. During Pavlovian learning, mice optimized their licking in two ways: they increased the fraction of time spent licking during the cue period from 8.98 ± 2.3% to 20.99 ± 3.1% (pre-learning vs post-learning, $p = 0.0078$, Two-tailed non-parametric Wilcoxon matched-pairs signed rank test), and they decreased non-contingent, 'spontaneous' licking during the ITI period from 8.73 ± 2.0% to 4.77 ± 1.2% (pre-learning vs post-learning, $p = 0.0273$, Two-tailed non-parametric Wilcoxon matched-pairs signed rank test) (**Fig. 2B-E**). Cue-evoked licking often ramped over the 3 second delay, indicating some knowledge of the temporal interval, though ramping on single trials was often variable within and across mice (**Fig. 2C**). A lick index was computed, reflecting the relative proportion of licking within the cue period relative to the ITI, to assess learning of the cue-reward association on each session (**Fig. 2D**).

Multi-phasic cholinergic interneuron firing and ACh release in the striatum are known to evolve with learning as Pavlovian cues become associated with reward [17,24,41]. ACh dips (or drops in interneuron firing rate) to reward-predictive cues become larger and broader from the naive to learned stages of Pavlovian conditioning[17,24]. These changes are inverse to, and may be modulated by, increases in cue-evoked DA release[19–21,42–45]. Consistent with previous findings, we observed, at some locations, multi-phasic, positive and negative changes in ACh release to conditioned light and tone cues after learning, consisting of a fast latency ('early') peak, a dip, and a second longer latency ('late') peak following the dip (**Fig. 2F-G**). Prior to learning, cues also evoked multi-phasic ACh release but at many fewer locations for all three components (**Fig. 2F-G**). Thus, Pavlovian associative learning was accompanied by the emergence of sequential, cue-evoked increases and decreases in ACh release across many striatum locations.

We then examined the spatial organization of learning related changes in the multi-phasic ACh release components across the striatum. Prior to learning, cue-evoked peaks and dips in ACh release were sparsely present across striatal locations, with isolated, large amplitude peaks in the tail of the striatum and dips scattered across the dorsal medial striatum (**Fig. 2H**). After learning, however, cue-evoked peaks and dips emerged across many locations in the anterior striatum with distinct spatial patterns, while posterior tail release remained largely stable (**Fig. 2F, H-I**). Across the dorsal striatum, prominent ACh dips emerged, which extended into the ventral striatum along the anterior wall (**Fig. 2F, H-I**). In the ventral striatum, a 'micro-organization' along the A/P axis was present, with the most anterior locations exhibiting a deepening of the cue-evoked dip in ACh, similar to dorsal, and more posterior locations an increase in the early peak with little or no dip (**Supplementary Fig. 2A**). Interestingly, early peaks in this region of the posterior ventral striatum occurred at longer latencies than peaks in other regions, indicating that they are likely distinct signals from dorsal peaks (**Supplementary Fig. 2B**). A majority of locations showed parallel learning related changes for both cues (**Supplementary Fig. 3**), but some locations showed cue selectivity (**Supplementary Fig. 3H**). Most notably, changes in the dip and early peak extended further into the posterior striatum for the tone than for the light, perhaps reflecting the organization of excitatory inputs to these regions[46,47] (**Supplementary Fig. 3F**). These results reveal diverse, spatially organized patterns of bi-directional, cue-evoked striatal ACh release dynamics that emerge during Pavlovian learning.

We next examined whether opposite changes were present at the time of water reward delivery as rewards became predicted with learning, as would be expected if ACh release were representing a 'prediction error' similar to that described for DA neuron firing and release [48–50]. Like the tone and light cues, water delivery was associated with spatially organized sequential peaks and dips in ACh release (**Fig. 1H and 3A-C**). The proportions and relative spatial organization of significant responses were similar between pre and post learning phases, but the relative amplitudes of the components changed (**Fig. 3B-D**). In contrast to dips at the conditioned cue, dips to the reward became smaller with learning across nearly the entire striatum, particularly dorsal (**Fig. 3D**). Changes to the initial and late peaks were more variable and less widespread, though there were opposing changes along the medial-lateral axis (**Fig. 3D**). Many of the same regions that showed an emergence of the cue-evoked dip also showed a diminished reward-related dip with learning for both cues, and there was a negative correlation between the magnitude of the reward dip and cue dip across sessions for both tone (Pearson's R^2= 0.21, p= $2.3x10^{-3}$) and light (R^2=0.12, p=$1.6x10^{-3}$) (**Fig. 3E**). These results are somewhat consistent with a partial reward prediction error signal, encoded by ACh dips. However, changes in the reward response with learning were slightly more widespread than for the cues (particularly the light), indicating that learning related changes to cues and rewards are partially decoupled and vary across regions (**Fig. 3D-E**). In addition, ACh peaks in the posterior ventral striatum, which increased with learning to the cue, did not reliably decrease with learning to the reward (**Fig. 3D**). Thus, changes in ACh dips to reward mirrors changes in ACh release to the cues, but only at particular striatum locations.

ACh signaling has been previously linked to changes in locomotion [23,38,51,52], so we asked whether the changes in cue and reward evoked ACh release after Pavlovian learning could be explained

by changes in treadmill velocity. A generalized linear model (GLM) was used to partially account for potential correlations with continuous linear velocity and acceleration (see Methods), but we conducted additional analyses to address potential movement contributions not captured by the GLM. For most locations, early peaks and dips for cues and rewards were highly consistent in timing and amplitude across trials within a given session, despite significant variability in the timing and magnitude of velocity changes (decelerations) at cue and reward (**Figs. 2, 3, and Supplementary Figs. 4 and 5**). ACh peaks and dips at only a small fraction of locations with significant response components were significantly correlated with trial by trial deceleration for cues (20/270, 16/224, and 2/180 locations for early peak, dip, and late peak of light cue respectively; 22/262, 9/176, and 3/76 locations for early peak, dip, and late peak of tone cue respectively; Pearson's correlation, $p<0.01$) and rewards (55/287, 65/276, and 14/229 locations for early peak, dip, and late peak respectively; Pearson's correlation, $p<0.01$). Thus, although changes in movement may modulate ACh release at some striatal locations, these correlations could not account for the patterns of cue and reward evoked release that we observed. Small artifacts at cue and reward were observed in null ACh sensor controls at a small fraction of sites, but they exhibited no spatial organization or learning related changes resembling the functional sensor, ruling out significant contributions of brain/implant motion or hemodynamic artifacts (**Supplementary Figs. 4H and 5H**).

Emergence of spatially organized ACh signals during extinction

Behavioral flexibility requires updating learned cue and action values as associated reward probabilities diminish, so we next asked how and where striatal ACh release reflects changes in learned predictive cue values (or incentive salience) during partial extinction learning. Following initial Pavlovian conditioning, the reward probability associated with one of the two cues was downshifted to 20% and the other to 80% (**Fig. 4A**). Cue-evoked anticipatory licking decreased over days specifically for the low probability cue, indicating successful cue-specific extinction learning (**Fig. 4B-C**). At many locations in the anterior dorsal striatum (aDS), we observed a dramatic emergence of long latency peaks in ACh release (late peaks), which were larger for the low probability (20%) cue but also occurred for the downshifted high probability (80%) cue (**Fig. 4D-F, J**; **Supplementary Figs. 6 and 7**). Elevations in late peaks following partial extinction were present for at least one fiber in 7 out of 8 mice for the light and 5 out of 8 mice for the tone (26/53, 49% and 17/53, 32% of all aDS fibers for light and tone respectively, see Methods). Importantly, the late peak increases were not simply a reversion to initial pre-learning levels, as very few late peaks were observed in these regions prior to learning (**Fig. 2H**). Dips that had emerged with initial learning in the dorsal striatum decreased during partial extinction (**Fig. 4F, H, J**; **Supplementary Figs. 6 and 7**). In contrast to the aDS, early peaks in ACh release in the posterior ventral striatum decreased to the low probability cue, a reversal of the changes occurring with initial learning (**Fig. 4G, J**). The amplitude of cue-evoked ACh peaks in the posterior tail did not change with partial extinction, similar to initial learning (**Fig. 4I, J**). The region-specific emergence of late peaks in aDS occurred many trials before the decreases in anticipatory spout licking to the 20% probability cue (light cue: median= 29 trials before licking decrease, STD=27, **Fig. 4K** and **L**; tone cue: median = 2 trials before licking decrease, STD =15, **Supplementary Figs. 6E and F**), indicating that these signals are positioned to drive learning related plasticity, perhaps

contributing to behavioral extinction (see Discussion). Release profiles were generally consistent between the tone and light cues with a few notable differences: late peaks were larger and more medial for the light relative to the tone (**Supplementary Fig. 6C, D, H** and **Supplementary Fig. 8**), and decreases in the amplitude of the dip were more prominent for the tone, particularly in the posterior striatum (**Supplementary Fig. 6D** and **Supplementary Fig. 8**). Changes in release following partial extinction could not be explained by velocity changes at cue onset, as peak deceleration amplitudes were not significantly different between the high and low probability cues or late learning prior to extinction (Kruskal-Wallis test, $p = 0.31$ for the light cue and $p = 0.97$ for the tone cue; for velocity changes, $p = 0.47$ for the light and $p = 0.99$ for the tone, **Fig. 4M** and **Supplementary Fig. 6G**). These data indicate that flexible downshifting of learned cue values is accompanied by bi-directional, region-specific changes in cue-evoked ACh release that precede behavioral changes. In the posterior ventral striatum, release reverts to pre-learning conditions when cues lose predictive value, whereas in the aDS, new elevations in ACh release emerge.

ACh release to unrewarded lick bouts changes with learning

Behavioral flexibility involves down-regulating actions that become less associated with reward, so we next tested how changes in ACh signaling reflect changes in the predictive value of spontaneous consummatory actions, independently of cues. During initial Pavlovian learning, mice decreased their spontaneous licking in the ITI period, indicating that they predict a higher probability of random reward early in learning relative to late after they have learned that reward delivery was more likely during cue periods than the ITI (**Fig. 2E**). Therefore, we asked whether ACh release was modulated around spontaneous ITI licks, and if so, whether lick related release changed from early to late learning (**Fig. 5**). Spontaneous lick bout initiations in the ITI pre-learning were associated with multi-phasic changes in ACh release at distinct latencies (**Fig. 5A-E**). Short latency increases in ACh began before the tongue contacted the lick spout, around lick initiation (**Fig. 5B, D**). Dips were present at intermediate latencies, while longer latency peaks occurred several hundred milliseconds after contact (**Fig. 5A-B**, **D**). The short and long latency peaks were present across largely distinct striatum locations (**Fig. 5E**). Short latency peaks were more prominent in the posterior striatum, while longer latency peaks were prominent in the anterior striatum (**Fig. 5A-B, E**). After learning, long latency peaks to ITI lick bout onsets decreased in magnitude in the aDS but increased within a small region of the dorsolateral striatum (**Fig. 5F**). Short latency peaks decreased at nearly all locations. Dips also became larger across the dorsal striatum (**Supplementary Fig. 9A**). None of these learning related changes could be explained by differences in locomotor deceleration magnitudes over learning, and they were not present in the null mice (**Supplementary Fig. 9B-K**). These results indicate that increases in ACh release across particular striatum sub-regions reflect changes in the prediction of reward associated with unrewarded spontaneous actions, signals which may facilitate flexible abandoning of unrewarded actions. Interestingly, late ACh peaks to unrewarded licks early in learning were present across the same regions of aDS where late peaks emerged to cues during extinction (**Fig. 4J**).

Localized ACh suppression impairs extinction of learned behavior

Our ACh measurements suggest that increases in transient ACh release in the aDS may contribute to downshifting the predictive value of cues and the expression of appetitive behaviors, such as licking, when reward probabilities change. We tested whether suppressing ACh release altered spontaneous and cue evoked spout licking during Pavlovian learning and partial extinction by virally expressing Tetanus toxin light chain (TelC) selectively in cholinergic interneurons in the aDS of ChAT-cre mice. TelC expression eliminates synaptic vesicle exocytosis in cholinergic interneurons, and consequently blocks ACh release[53] (**Fig. 6A-B**). ACh suppression did not affect the mice ability to learn the association between the cue and reward, as measured by changes in the frequency of licking in the cue period, relative to the ITI (Two-way repeated measure ANOVA, $p = 0.02$, **Fig. 6C**). This is expected, given that the primary changes in ACh release in aDS during initial learning were in the dips, which may be permissive for learning[25,54–56] (**Fig. 2I**). During cue specific extinction, control mice exhibited a down-regulation of anticipatory licking relative to the post-learning phase, as expressed by the significant decrease in the fraction of time spent licking during cue period ($p = 0.04$, Two-way repeated measures ANOVA, **Fig. 6C** and **E**). However, TelC expressing mice did not show a significant down-regulation of cue period licking relative to the learned phase ($p = 0.93$, Two-way repeated measures ANOVA, **Fig. 6C** and **E**). TelC mice also had higher, but not statistically significant, licking rates during the ITI and cue periods than the control mice at all phases of learning (**Fig. 6C** and **D**). These data indicate that elevations in ACh release in the aDS contribute to the extinction of learned Pavlovian responses following a downshift in cue associated reward probabilities.

Dopamine and glutamate release dynamics during extinction indicate a local mechanism for ACh increases

We next investigated potential contributors to the emergence of the downshifted cue and spontaneous lick related ACh increases in the aDS. Dips in DA neuron firing and release have been measured in response to low value cues and unpredicted reward omissions (i.e. negative prediction errors) and have been proposed to contribute to extinction learning[48,57,58]. If present, these DA dips may disinhibit CINs through reduced tonic activation of inhibitory D2 receptors[19,53,59,60]. We tested whether DA signals in the aDS were consistent with a potential role in facilitating the observed elevations in ACh signaling. We optically measured DA release dynamics with dLight1.3b[61] in the same regions of the aDS where we observed the ACh increases (**Fig. 7A**). Consistent with prior work, DA release rapidly increased to conditioned cues following learning (**Fig 7B**). Following partial extinction (**Fig. 4A**), dips below baseline emerged after short latency peaks, selectively for the low probability cue (**Fig. 7B-D**). The timing of DA dips (median to trough = 0.63s) preceded the timing of the ACh long latency peaks (median = 0.8s) in the same regions (**Fig. 7E**, $p = 0.02$, Two-tailed rank-sum test), and dips emerged within only a few trials after the probability shift, well before the change in cue-evoked licking (median = 27 trials prior to lick change, STD=23.4 for light, median = 22 trials, STD=42.8 for tone, **Fig. 7F**). We then asked whether DA release for spontaneous ITI licks also exhibited learning related dips reflecting shifts in reward prediction with learning. DA release in the aDS increased for spontaneous licks, rising just prior to spout contact. Early in Pavlovian learning, peaks were followed by dips below

baseline, but these dips largely disappeared after learning ($p = 2.6x10^{-3}$, two-tailed Wilcoxon rank-sum test), mirroring opposite changes in ACh release as ITI licks became less frequent **(Fig. 7H-J**). These results indicate that DA signals in the aDS exhibit emerging dips to cues after partial extinction and spontaneous ITI licks prior to learning which precede ACh elevations in the same striatal region (DA median latency to dip = 0.33s vs ACh median latency to long latency peak= 0.44s, $p= 3.33x10^{-6}$, two-tailed Wilcoxon rank-sum test, **Fig. 7K**).

Next, we asked whether emerging ACh increases to cues after partial extinction may develop as a consequence of increases in excitatory glutamatergic drive. To test this possibility, we expressed the genetically encoded glutamate sensor iGluSnFr[62] selectively in aDS cholinergic interneurons of ChAT-cre mice to measure rapid changes in glutamate release (GluCIN) during partial extinction (**Fig. 8**). Interestingly, GluCIN release in the aDS to the light stimulus following Pavlovian learning was bi-phasic, consisting of an initial fast latency increase followed by a slower latency increase (**Fig. 8E**). Consistent with the lack of strong ACh release in the aDS to the conditioned tone (**Supplementary Fig. 6B-D**), no rapid changes in aDS GluCIN release were present for the tone, despite similar pre-licking behavior and velocity changes (**Fig. 8E**). This observation ruled out potential contributions of movement changes to the fast bi-phasic profile in the aDS GluCIN signal. After the downshift in reward probability associated with the light cue (from 100 to 20%), there was no significant change in this bi-phasic profile (**Fig. 8E**, fast latency peak: $p = 0.96$, two-tailed Wilcoxon rank-sum test; slow latency peak: $p = 0.34$, two-tailed Wilcoxon rank-sum test) indicating that an increase in glutamate release onto the CIN was not responsible for the increase of aDS ACh release during partial extinction. Despite this, the slower component of the GluCIN release aligned with the timing of the ACh late peak increase following extinction (median latency to Glu slow peak = 0.64s, vs median latency to ACh late peak = 0.8s), indicating that the influence of this input may be 'unmasked' to drive the ACh increase, perhaps by disinhibition through emergent DA dips (see Discussion).

Discussion

We applied high density arrays of small diameter optical fibers to map the spatiotemporal dynamics of rapid changes in ACh release across the striatum during Pavlovian learning and partial extinction in behaving mice. Our measurements revealed regional differences in how ACh release to cues, rewards, and consummatory actions evolves across timescales ranging from 100s of milliseconds to days as Pavlovian associations are learned then broken. During initial learning, we observed region-specific patterns of ACh dips to cues and rewards that changed in partial accordance with positive prediction error encoding, though in opposite directions to widely described striatal DA release[48]. The changes in multi-phasic ACh signaling we observed during initial learning across the dorsal and ventral anterior striatum are largely consistent with previous observations in rodents and non-human primates[15,17,19,23,24]. Mechanistic work has suggested that ACh dips may be permissive for enabling DA elevations to drive potentiation of excitatory synapses onto D1 receptor expressing projection neurons during associative learning[26,27,63–65]. Notably, however, changes in ACh dips and peaks during initial learning were not present across all striatum regions. For example, in the posterior, but not anterior, ventral striatum, early peaks

in ACh signaling increased to cues and decreased to rewards, with few changes in dips, a signaling pattern similar to DA[23,66] (**Fig. 2F-I** and **Supplementary Fig. 2**). In the posterior tail, ACh release strongly increased to cues but did not change with learning (**Fig. 2F-I**). These regional differences in ACh signaling are likely driven, at least in part, by the topography of glutamatergic and neuromodulatory (e.g. dopaminergic) inputs to striatal CINs[46,47,59], a notion supported by the differences in signal topography to light and tone cues (**Supplementary Fig. 3**). Thus, a 'permissive' role for ACh dips in learning may be region-selective and depend on specific task parameters (e.g. cue sensory modality).

Following a down-shift of learned cue values during partial extinction learning, cue-evoked ACh release concentrated in a region of the aDS did not simply revert back to pre-learning levels (as in other regions) but acquired new emergent peaks which preceded, over trials, a down-shift in anticipatory licking, selectively for low probability cues (**Fig. 4**). Prior to initial learning, these same aDS regions displayed ACh release peaks following unrewarded, spontaneous lick initiations, which diminished with learning, consistent with signaling of unpredicted reward omissions associated with consummatory actions (**Fig. 5**). Collectively, these results suggest that transient ACh release in the anterior striatum may actively downregulate the relative motivational salience (or value) of learned conditioned cues and the expression of specific actions (e.g. licking) as they become decoupled with reward. During initial Pavlovian learning, ACh peaks following spontaneous, unrewarded licks may downregulate the inappropriate expression of non-contingent licking behavior outside of the cue period. Similarly, during cue-specific extinction, emerging ACh peaks to the low probability cue may decrease learned anticipatory licking by downshifting the cue's learned predictive value. Consistent with this hypothesis, we found that persistent inhibition of ACh release from aDS CINs prevented the down-regulation of licking during partial extinction (**Fig. 6**). While we did not observe a significant difference in ITI licking following inhibition of aDS ACh release, ITI lick signals during learning were more widespread and consisted of topographically organized early and late components, which may represent the relative action value and/or unexpected reward omission (result of the lick) respectively (**Fig. 5E-F**). Our data suggest that phasic increases in ACh signaling in the aDS may facilitate behavioral flexibility by erasing representations of learned cue values or context-dependent action contingencies when their associated reward probability is diminished. Although task-relevant ACh dynamics to cues and actions were prominent and highly concentrated in the aDS, shifts were also observed across other regions, which may also play a role in extinction learning. Indeed, another study observed a correlation between ACh transients in the medial nucleus accumbens shell and cocaine place preference extinction[9]. Additional studies will clarify the specific ACh dynamics associated with different forms of behavioral flexibility across striatal regions.

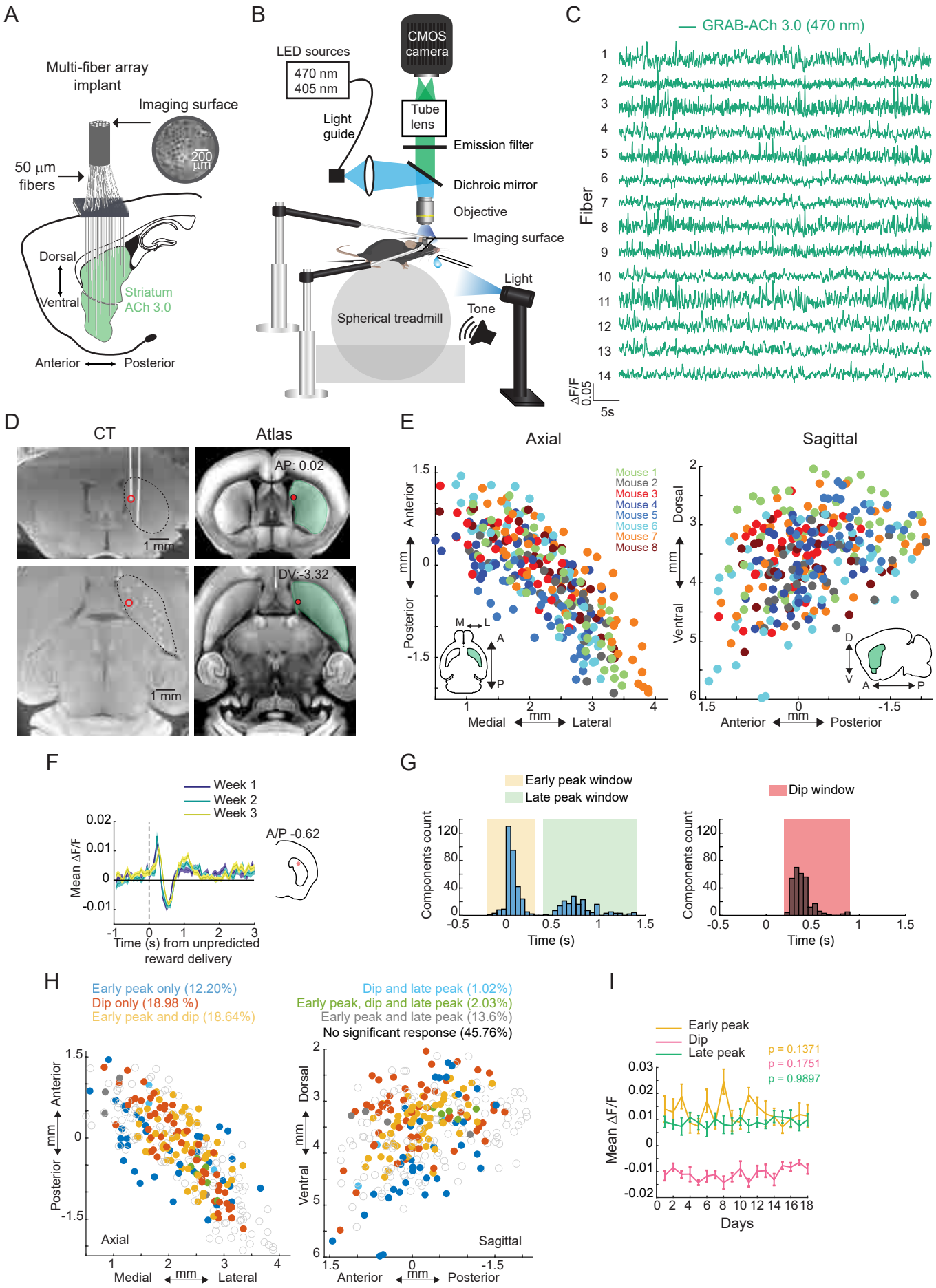
DA has been long known to modulate excitability of cholinergic interneurons in healthy and pathological states[67], so we investigated whether DA dynamics are positioned to facilitate the emergence of transient ACh release in the aDS. We found that DA release dipped below baseline following unrewarded ITI licks early in learning and following low probability cues after partial extinction (**Fig. 7**). These dynamics emerged prior to behavioral changes and preceded (on average), the late ACh peaks in the same region by approximately 100-200ms (**Fig. 7E-F**). Optogenetic inhibition of DA release is capable of rapidly (<200ms) elevating ACh release *in-vivo*

via reduced binding of baseline (tonic) DA to inhibitory D2 receptors[53]. Thus, the dips in DA release we observed may facilitate (or drive) the emergence of aDS ACh increases to ITI licks and extinction cues. Supporting the notion of an intrastriatal mechanism, we found that cue-evoked glutamate release onto CINs did not change during partial extinction but did exhibit two distinct peaks, which aligned with fast and slow latency peaks in cue-evoked ACh release (**Fig. 8**). The late component therefore may not drive strong ACh release in the learned phase but may be 'unmasked' to influence release following the emergence of DA dips during extinction. Future studies specifically targeting D2 receptor signaling on aDS CINs will be needed to test this possible mechanism.

How might transient ACh signaling during extinction act to erase plasticity for previously valued cues or actions? Studies *in-vitro* have found evidence that M4 muscarinic ACh receptors, expressed predominantly on direct pathway spiny projection neurons (dSPNs), promote long term synaptic depression[68,69]. One possibility is that the increases in cue evoked ACh during extinction drive depression of previously potentiated synapses onto dSPNs via M4 receptors. Potentiation of dSPN synapses is believed to drive initial associative learning through DA binding to D1 receptors, which occurs during phasic elevations in DA in response to unpredicted rewards and reward-associated cues (e.g. positive reward prediction errors)[27,48,63]. Dips in DA release were observed in our task to downshifted cues and unrewarded ITI licks, and DA dips, encoding negative prediction errors, have been proposed to contribute to extinction. However, it is unclear how DA dips alone could influence the strength of previously potentiated dSPN synapses because low affinity D1 receptors are not strongly activated by basal DA levels and therefore are not sensitive to DA dips[70]. In contrast, CINs (and D2 expressing projection neurons, iSPNs), which express the high affinity D2 DA receptor, are sensitive to DA dips, and ACh levels in the striatum increase rapidly when DA release is inhibited[53,70]. Thus, disinhibition of CINs may provide a mechanism to convert negative prediction errors encoded by DA dips into ACh elevations that can drive de-potentiation of dSPN synapses (or potentiation of iSPNs) to erase learned associations.

Figures

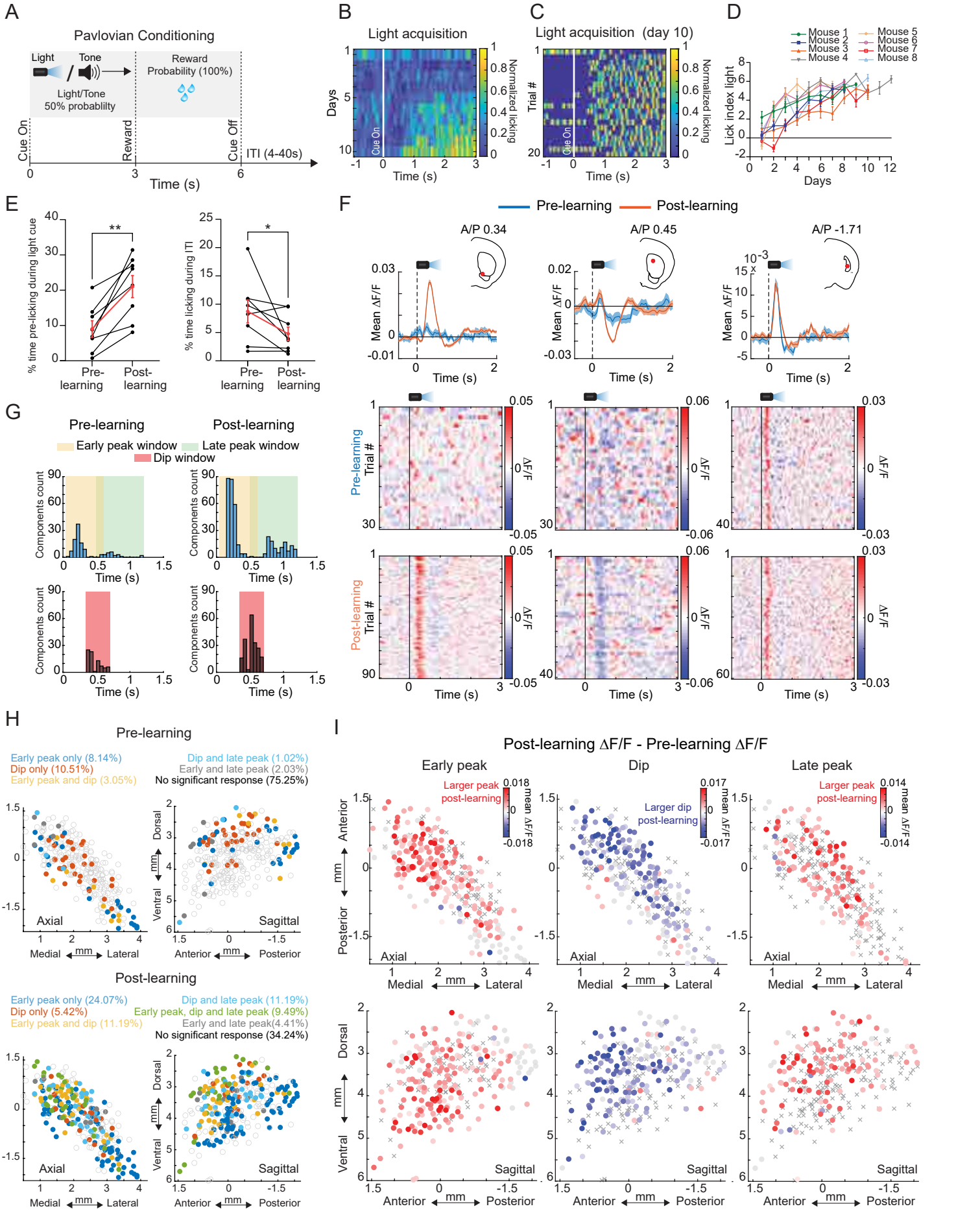
FIGURE 1


A
Multi-fiber array implant
Imaging surface
200 µm
50 µm fibers
Dorsal
Ventral
Striatum ACh 3.0
Anterior
Posterior
B
CMOS camera
LED sources
470 nm
405 nm
Light guide
Tube lens
Emission filter
Dichroic mirror
Objective
Imaging surface
Light
Tone
Spherical treadmill
C
GRAB-ACh 3.0 (470 nm)
Fiber
ΔF/F
0.05
5s
D
CT
Atlas
AP: 0.02
1 mm
DV:-3.32
E
Axial
Sagittal
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Mouse 6
Mouse 7
Mouse 8
Anterior
Posterior
Medial
Lateral
Dorsal
Ventral
mm
F
Week 1
Week 2
Week 3
Mean ΔF/F
A/P -0.62
Time (s) from unpredicted reward delivery
G
Early peak window
Late peak window
Dip window
Components count
Time (s)
H
Early peak only (12.20%)
Dip only (18.98 %)
Early peak and dip (18.64%)
Dip and late peak (1.02%)
Early peak, dip and late peak (2.03%)
Early peak and late peak (13.6%)
No significant response (45.76%)
Axial
Sagittal
I
Early peak
Dip
Late peak
p = 0.1371
p = 0.1751
p = 0.9897
Mean ΔF/F
Days

<u>Figure 1:</u> Chronic striatum-wide measurements of rapid ACh release dynamics in head-fixed, behaving mice

A. Schematic of the fiber array approach for measuring ACh release across the striatum. **B.** Spherical treadmill and imaging set up for head-fixed mice. **C.** ACh3.0 fluorescence traces ($\Delta F/F$) measured from 14 example fibers in the striatum of a single mouse. **D.** Left: Post-mortem micro-CT scan images in the coronal plane (top) and axial plane (bottom) from a representative mouse. Fibers appear in white. Right: Images from the Allen Brain Common Coordinate Framework Atlas corresponding to the CT planes on left. Red circles indicate the position of an automatically localized fiber tip. **E.** Locations of all fibers used for ACh measurements in the axial (left) and sagittal (right) planes. Each dot is the location (relative to bregma) of a single fiber and the color indicates the mouse identity (n = 8 mice). **F.** Mean $\Delta F/F$ aligned to unpredicted water reward delivery for a single fiber for single sessions (n=8 reward deliveries/session) at 1-3 weeks from the first day of Pavlovian training (Figure 2). Inset indicates fiber location in the coronal plane. Shaded regions, S.E.M. **G.** Histograms of the number of fibers expressing a significant ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods) early peak, dip, and late peak response component to unpredicted reward delivery. Time windows used to define each component are indicated by shaded regions. **H.** Maps as in (E) showing the presence of each combination of signal components to unpredicted reward delivery for each fiber. Empty circles indicate no significant response. **I.** The mean peak (early/late) or dip $\Delta F/F$ for unpredicted reward deliveries across 18 consecutive sessions of training for the same example fiber in (F). Values are the mean ± S.E.M. For each component, the mean of each day was compared to the mean of every other day using one-way ANOVA.

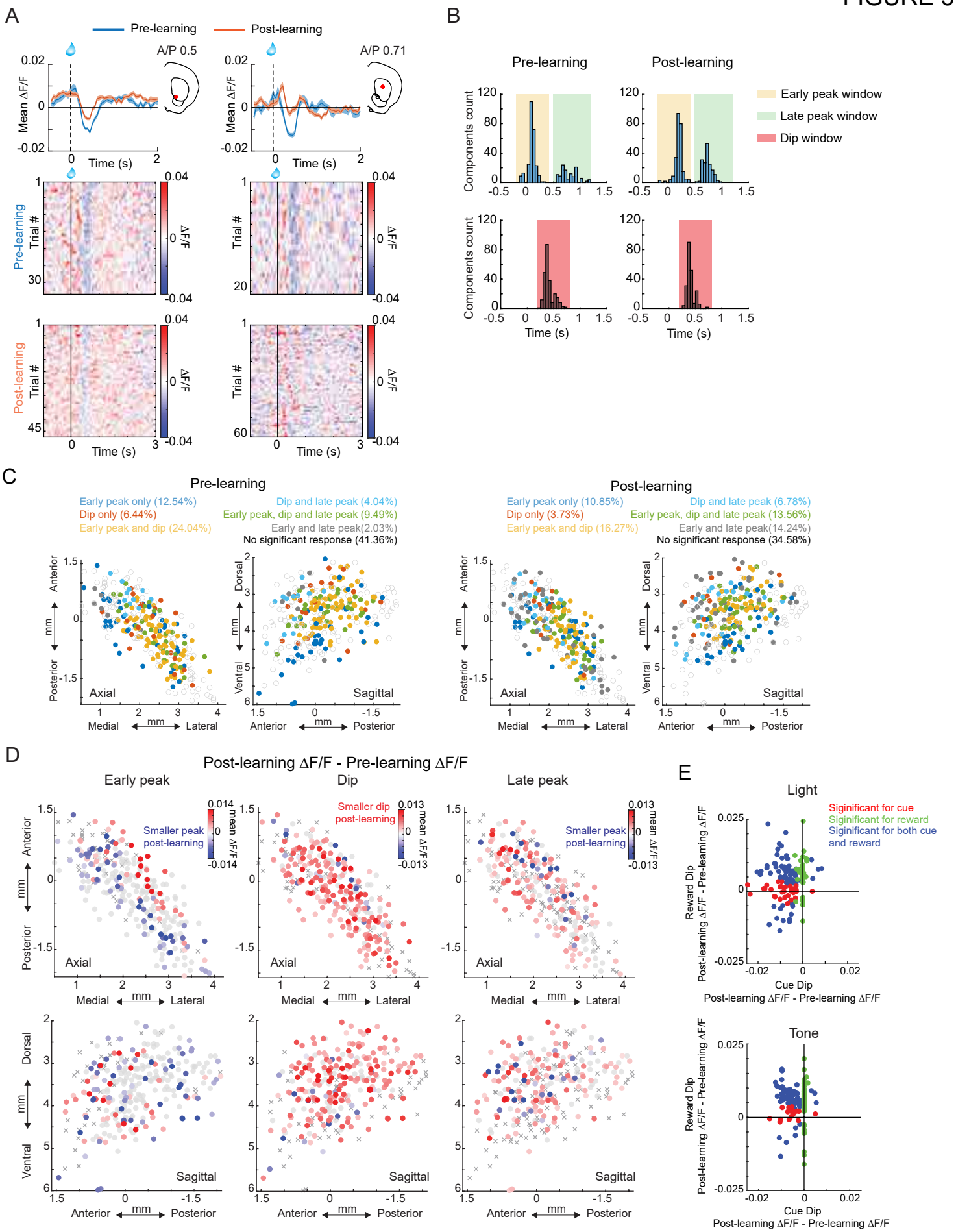
FIGURE 2


A
Pavlovian Conditioning
Light
Tone
Reward Probability (100%)
Light/Tone 50% probablity
Cue On
Reward
Cue Off
ITI (4-40s)
Time (s)
B
Light acquisition
Days
Cue On
Time (s)
Normalized licking
C
Light acquisition (day 10)
Trial #
Cue On
Time (s)
Normalized licking
D
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Mouse 6
Mouse 7
Mouse 8
Lick index light
Days
E
% time pre-licking during light cue
Pre-learning
Post-learning
% time licking during ITI
F
Pre-learning
Post-learning
A/P 0.34
A/P 0.45
A/P -1.71
Mean ΔF/F
Time (s)
Pre-learning
Trial #
Post-learning
Trial #
ΔF/F
G
Pre-learning
Post-learning
Early peak window
Late peak window
Dip window
Components count
Time (s)
H
Pre-learning
Early peak only (8.14%)
Dip only (10.51%)
Early peak and dip (3.05%)
Dip and late peak (1.02%)
Early and late peak (2.03%)
No significant response (75.25%)
Axial
Sagittal
Medial
Lateral
Anterior
Posterior
Dorsal
Ventral
mm
Post-learning
Early peak only (24.07%)
Dip only (5.42%)
Early peak and dip (11.19%)
Dip and late peak (11.19%)
Early peak, dip and late peak (9.49%)
Early and late peak(4.41%)
No significant response (34.24%)
I
Post-learning ΔF/F - Pre-learning ΔF/F
Early peak
Dip
Late peak
Larger peak post-learning
Larger dip post-learning
mean ΔF/F
Axial
Sagittal
Medial
Lateral
Anterior
Posterior
Dorsal
Ventral
mm

<u>Figure 2:</u> Region specific, bi-directional changes in multi-phasic ACh release to conditioned light cues during Pavlovian learning

A. Schematic of dual-cue delay Pavlovian conditioning task. On each trial, either a light or tone cue was presented (4-40s ITI) followed by a water reward delivery after 3s. The cue was left on until 3s after reward delivery. **B.** Normalized mean lick spout contacts aligned to light cue onset for 10 consecutive sessions from the start of training for one representative mouse. **C.** Normalized mean lick spout contacts as in (B) for 20 trials in a single session (day 10) after learning. **D.** Mean lick index on light cue trials for each day after the start of Pavlovian training for each mouse. Error bars, S.E.M. **E.** Mean percent time spent licking during the 3s light cue interval prior to reward (left) and during the ITI (right) for each individual mouse for sessions pre and post learning (phases determined by the overall lick index, see STAR Methods). Red points are means (± S.E.M.) across all mice (n = 8), significant difference between pre- and post-learning assessed by a two-tailed non-parametric Wilcoxon matched-pairs signed rank test (*$p<0.05$, **$p<0.01$). **F.** Top: Mean $\Delta F/F$ for 3 representative fibers aligned to the light cue onset for trials pre (blue) and post (orange) learning. Shaded regions, S.E.M. Red dots in insets indicate the fiber locations in the coronal plane. Bottom: Light-cue-aligned $\Delta F/F$ for all trials included in the triggered averages at top. **G.** Number of fibers expressing a significant ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods) early peak, dip, and late peak response component to light cue onset pre (left) and post (right) learning. **H.** Maps (axial, left; sagittal, right) indicating, by color, the presence of each combination of significant large amplitude signal components (see Methods) to light cue onset for each fiber (circle) pre (top) and post (bottom) learning. Empty circles are fibers with no significant response. **I.** Maps as in (H) showing the difference in mean light-cue evoked $\Delta F/F$ with learning (post – pre) for the three signal components. Gray circles, no significant difference; Xs, no significant component. Significance of each component for each fiber was assessed using the two-tailed Wilcoxon rank-sum test, ($p < 0.01$).

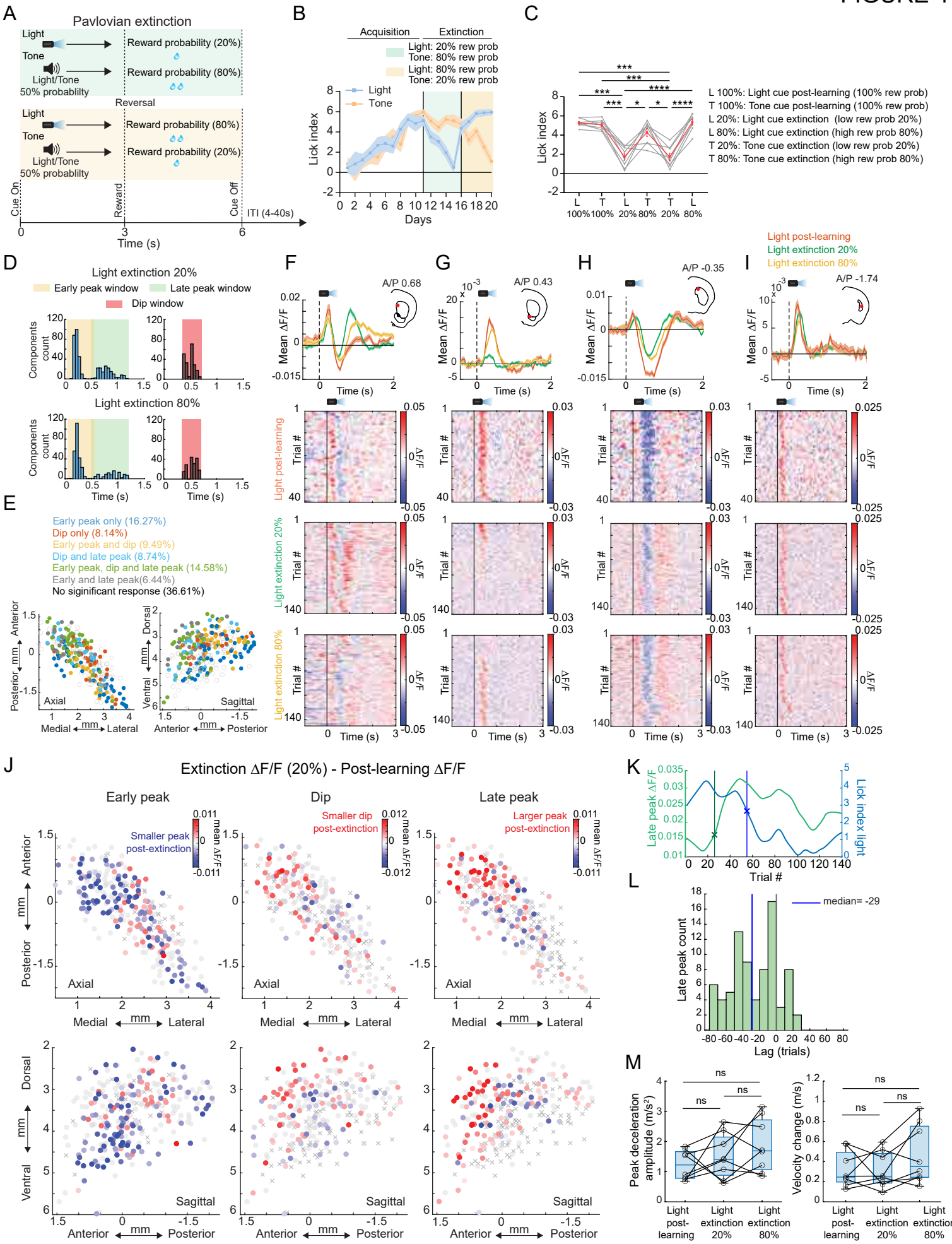
FIGURE 3


A
Pre-learning
Post-learning
A/P 0.5
A/P 0.71
Mean ΔF/F
Time (s)
Trial #
ΔF/F
B
Pre-learning
Post-learning
Components count
Early peak window
Late peak window
Dip window
Time (s)
C
Pre-learning
Early peak only (12.54%)
Dip only (6.44%)
Early peak and dip (24.04%)
Dip and late peak (4.04%)
Early peak, dip and late peak (9.49%)
Early and late peak(2.03%)
No significant response (41.36%)
Post-learning
Early peak only (10.85%)
Dip only (3.73%)
Early peak and dip (16.27%)
Dip and late peak (6.78%)
Early peak, dip and late peak (13.56%)
Early and late peak(14.24%)
No significant response (34.58%)
Anterior
Posterior
Dorsal
Ventral
Medial
Lateral
Axial
Sagittal
mm
D
Post-learning ΔF/F - Pre-learning ΔF/F
Early peak
Dip
Late peak
Smaller peak post-learning
Smaller dip post-learning
mean ΔF/F
E
Light
Tone
Siginificant for cue
Siginificant for reward
Siginificant for both cue and reward
Reward Dip
Post-learning ΔF/F - Pre-learning ΔF/F
Cue Dip
Post-learning ΔF/F - Pre-learning ΔF/F

<u>Figure 3:</u> Widespread changes in ACh dips to rewards during Pavlovian learning

A. Top: Mean ∆F/F for trials pre- and post-learning aligned to reward consumption for two representative fibers in two mice for trials pre (blue) and post (orange) learning. Shaded regions, S.E.M. Red dots in insets indicate fiber locations in the coronal plane. Bottom: Reward-aligned ∆F/F for all trials included in the triggered averages at top. **B.** Histogram of the number of fibers expressing a significant early peak, dip, and late peak response component to light cue onset pre (left) and post (right) learning ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods). **C.** Maps (axial, left; sagittal, right) indicating the presence of each combination of signal components to reward consumption for each fiber (circle) pre (left) and post (right) learning. Empty circles are fibers with no significant response. **D.** Maps showing the difference in mean reward-evoked ∆F/F with learning (post – pre) for the three signal components. Gray circles, no significant difference; Xs, no significant component. The significance of each component for each fiber was calculated using a two-tailed Wilcoxon rank-sum test ($p < 0.01$). **E.** Difference in mean reward-evoked ∆F/F with learning (post – pre) vs the difference in mean light (top) or tone (bottom) cue-evoked ∆F/F dip with learning. Each dot is one fiber in the anterior striatum (AP >=-0.5mm) and color indicates whether the ∆F/F difference with learning is significant (two-tailed Wilcoxon rank-sum test, $p<0.05$) for the reward, cue, or both.

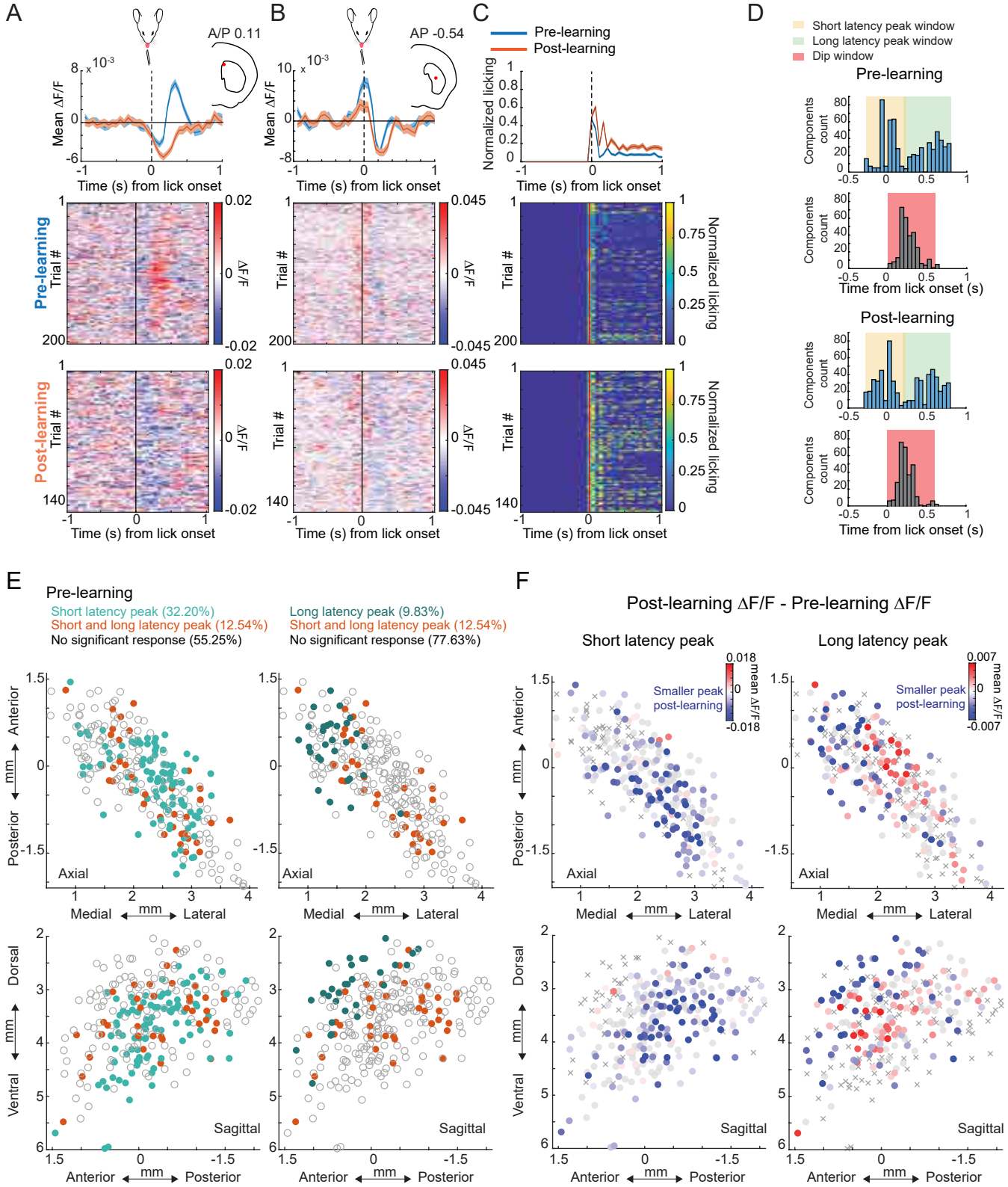
FIGURE 4


A
Pavlovian extinction
Light
Reward probability (20%)
Tone
Reward probability (80%)
Light/Tone
50% probablilty
Reversal
Light
Reward probability (80%)
Tone
Reward probability (20%)
Light/Tone
50% probablilty
Cue On
Reward
Cue Off
ITI (4-40s)
Time (s)
B
Acquisition
Extinction
Light: 20% rew prob
Tone: 80% rew prob
Light: 80% rew prob
Tone: 20% rew prob
Light
Tone
Lick index
Days
C
Lick index
L 100%: Light cue post-learning (100% rew prob)
T 100%: Tone cue post-learning (100% rew prob)
L 20%: Light cue extinction (low rew prob 20%)
L 80%: Light cue extinction (high rew prob 80%)
T 20%: Tone cue extinction (low rew prob 20%)
T 80%: Tone cue extinction (high rew prob 80%)
D
Light extinction 20%
Early peak window
Late peak window
Dip window
Components count
Light extinction 80%
Time (s)
E
Early peak only (16.27%)
Dip only (8.14%)
Early peak and dip (9.49%)
Dip and late peak (8.74%)
Early peak, dip and late peak (14.58%)
Early and late peak(6.44%)
No siginificant response (36.61%)
Axial
Sagittal
Medial
Lateral
Anterior
Posterior
Dorsal
Ventral
F
G
H
I
A/P 0.68
A/P 0.43
A/P -0.35
A/P -1.74
Light post-learning
Light extinction 20%
Light extinction 80%
Mean ΔF/F
Trial #
Time (s)
J
Extinction ΔF/F (20%) - Post-learning ΔF/F
Early peak
Dip
Late peak
Smaller peak post-extinction
Smaller dip post-extinction
Larger peak post-extinction
mean ΔF/F
Axial
Sagittal
Medial
Lateral
Anterior
Posterior
Dorsal
Ventral
K
Late peak ΔF/F
Lick index light
Trial #
L
median= -29
Late peak count
Lag (trials)
M
Peak deceleration amplitude (m/s²)
Velocity change (m/s)
ns
Light post-learning
Light extinction 20%
Light extinction 80%

<u>Figure 4:</u> Emergence of light cue-evoked increases in striatal ACh release in the anterior striatum during partial cue-specific extinction precedes behavioral changes

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.** lick index 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. *: $p < 0.05$, ***: $p < 0.001$, ****: $p < 0.0001$. **D.** Histogram of the number of fibers expressing a significant ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods) early peak, dip, and late peak response component to light cue onset associated with 80% or 20% reward probability. **E.** Maps (axial, left; sagittal, right) indicating, by color, the presence of each combination of signal components to 20% reward probability light cue onset for each fiber (circle). Empty circles are fibers with no significant response. **F.** Top: Mean $\Delta F/F$ for a representative fiber aligned to light cue onset for trials where light was associated with 100% (post-learning), 80%, and 20% (extinction) probabilities. Shaded regions, S.E.M. Red dot in the inset indicates the fiber location. Bottom: Light-cue-aligned $\Delta F/F$ for all trials included in the triggered averages at top. **G-I.** Same as (F) for other representative fibers. **J.** Maps showing the difference in mean light cue-evoked $\Delta F/F$ between the post-learning (100%) and 20% reward probability phases for the three signal components. Each circle is one fiber. Gray circles, no significant difference; Xs, no significant component. Significance differences for each component for each fiber was assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$). **K.** 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. **L.** 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. **M.** Box and whisker plots showing the amplitudes of the peak decelerations and velocity changes across all mice (n=8 mice) following light cue onsets. n.s., not significant (Kruskal-Wallis test, $p > 0.05$). Each datapoint is the mean for one mouse.

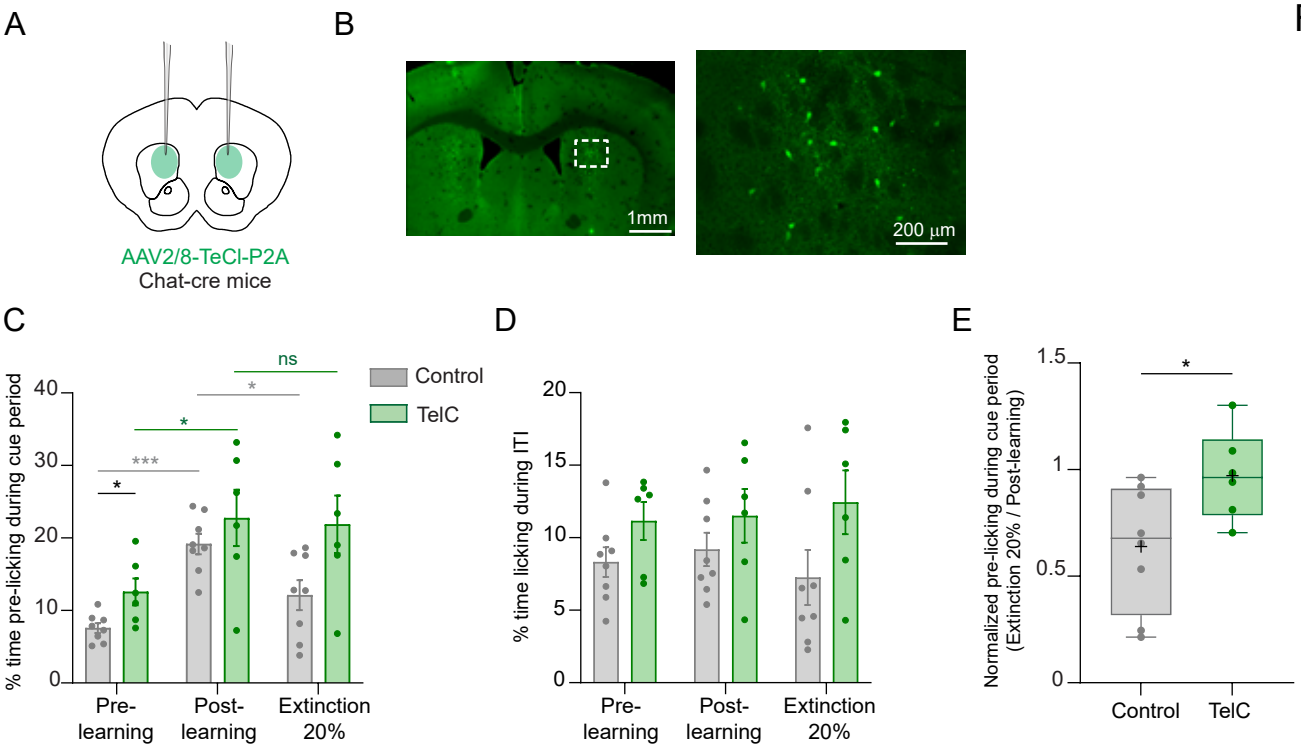
FIGURE 5


A
A/P 0.11
Mean ΔF/F
Time (s) from lick onset
Pre-learning
Post-learning
Trial #
ΔF/F
B
AP -0.54
C
Pre-learning
Post-learning
Normalized licking
D
Short latency peak window
Long latency peak window
Dip window
Pre-learning
Components count
Time from lick onset (s)
Post-learning
E
Pre-learning
Short latency peak (32.20%)
Short and long latency peak (12.54%)
No significant response (55.25%)
Long latency peak (9.83%)
Short and long latency peak (12.54%)
No significant response (77.63%)
Anterior
Posterior
mm
Axial
Medial
Lateral
Dorsal
Ventral
Sagittal
F
Post-learning ΔF/F - Pre-learning ΔF/F
Short latency peak
Long latency peak
Smaller peak post-learning
mean ΔF/F

<u>Figure 5:</u> Changes in region-specific ACh release to spontaneous unrewarded, lick bouts in the intertrial interval period during learning

A. Mean ∆F/F for a representative fiber aligned to the onset of spontaneous, unrewarded spout licking bouts in the ITI on trials pre (blue) and post (orange) initial Pavlovian learning (see Figure 2). Shaded regions, S.E.M. Red dot in the inset indicates the fiber location in the coronal plane. Bottom: Lick-onset-aligned ∆F/F for all onsets included in the triggered averages at top. **B.** Same as (A) for a different fiber in the same mouse. **C.** Top: Normalized mean lick spout contacts aligned to ITI lick bouts onsets pre and post learning for the onsets in (A-B). Shaded region, S.E.M. Bottom: Normalized bout-onset-aligned lick counts for all onsets included in (A-B). **D.** Histogram of the number of fibers expressing significant ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods) early peak, dip, and late peak response components to ITI lick bout onsets pre (top) and post (bottom) learning. **E.** Maps (axial, top; sagittal, bottom) indicating the presence of each combination of signal components to ITI lick bout onsets for each fiber (circle) pre learning. Empty circles are fibers with no significant response. **F.** Maps showing the difference in mean ITI lick bout-evoked ∆F/F between pre- and post-learning for the early and late peaks. Each circle is one fiber. Gray circles, no significant difference; Xs, no significant component. Significance differences for each component for each fiber were assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$).

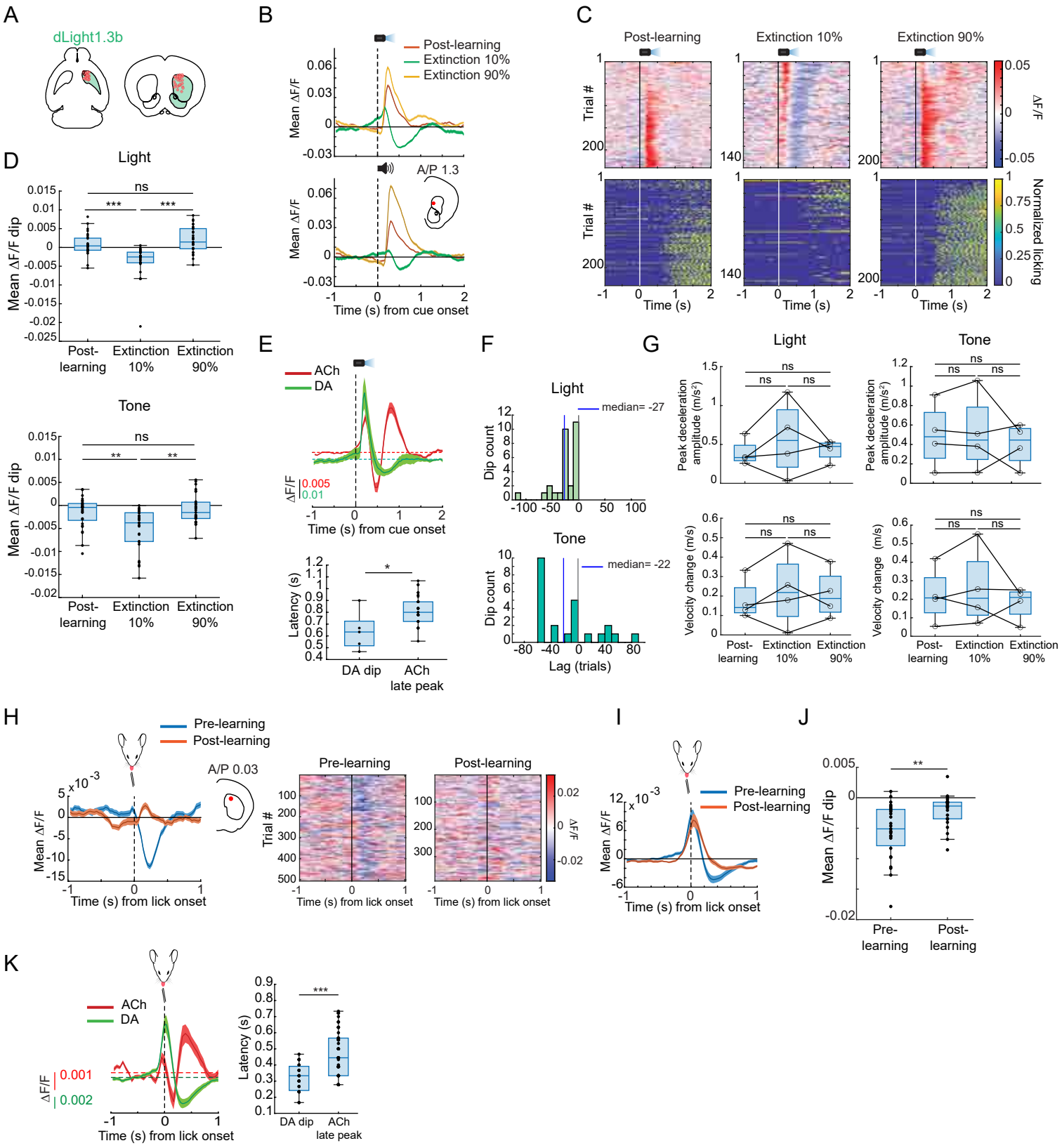
FIGURE 6


A
AAV2/8-TeCl-P2A
Chat-cre mice
B
1mm
200 µm
C
% time pre-licking during cue period
Control
TelC
ns
*
*
***
*
Pre-learning
Post-learning
Extinction 20%
D
% time licking during ITI
Pre-learning
Post-learning
Extinction 20%
E
Normalized pre-licking during cue period (Extinction 20% / Post-learning)
*
Control
TelC

<u>Figure 6:</u> Inactivation of acetylcholine 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 cholinergic interneurons. **C.** Mean (± 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 as multiple comparison within and between TelC and control mice for all phases of learning). **D.** same as (C) for licking during ITI periods. **E.** Box plots showing the ratio of mean cue period licking for 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-values were calculated with a Mann-Whitney test (two-tailed) to compare the two groups. *p<0.05.

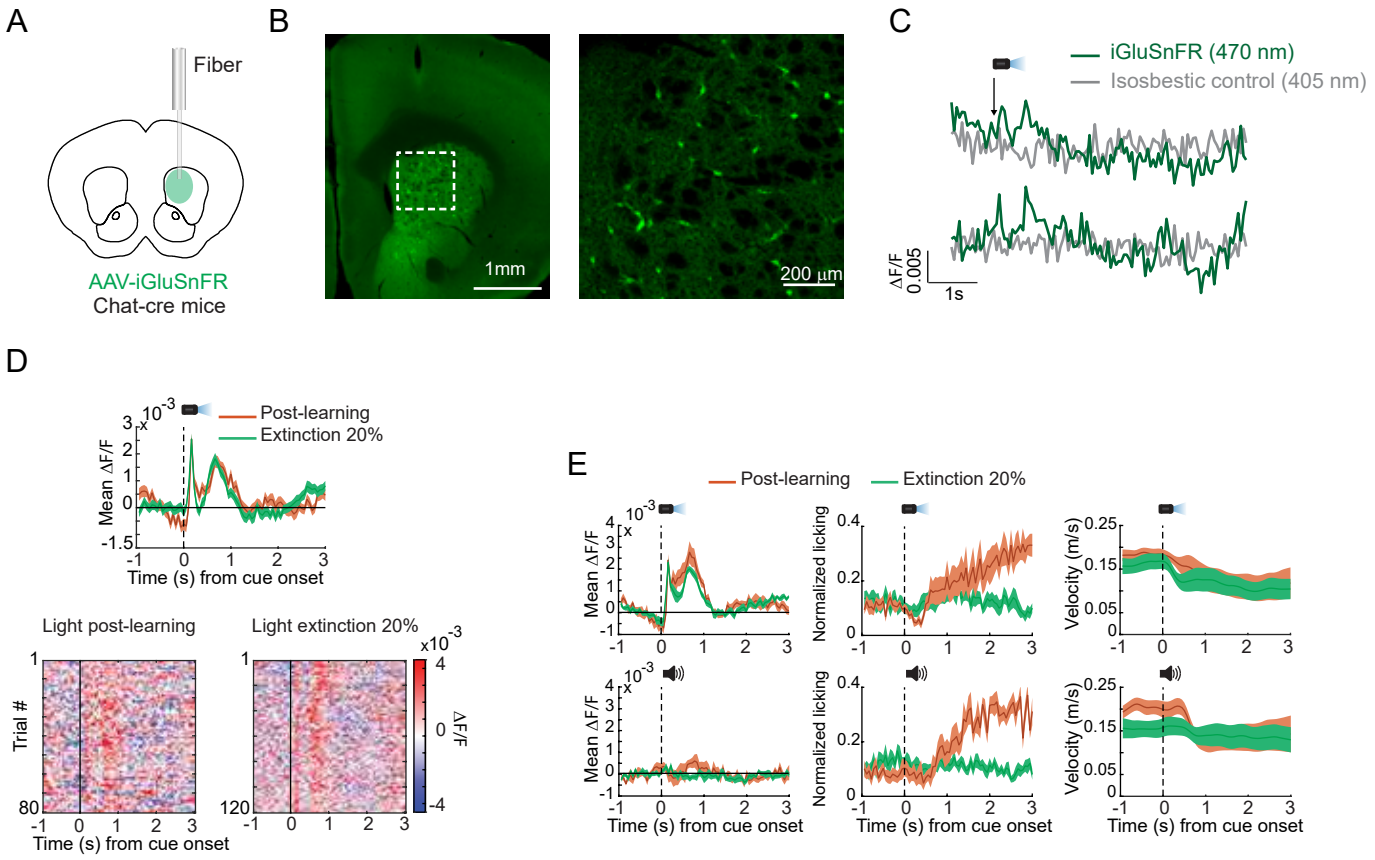
FIGURE 7


A
dLight1.3b
B
Post-learning
Extinction 10%
Extinction 90%
Mean ΔF/F
A/P 1.3
Time (s) from cue onset
C
Post-learning
Extinction 10%
Extinction 90%
Trial #
ΔF/F
Normalized licking
Time (s)
D
Light
Tone
Mean ΔF/F dip
ns
***
**
Post-learning
Extinction 10%
Extinction 90%
E
ACh
DA
ΔF/F
Time (s) from cue onset
Latency (s)
DA dip
ACh late peak
F
Light
Tone
Dip count
median= -27
median= -22
Lag (trials)
G
Light
Tone
Peak deceleration amplitude (m/s²)
Velocity change (m/s)
ns
Post-learning
Extinction 10%
Extinction 90%
H
Pre-learning
Post-learning
A/P 0.03
Mean ΔF/F
Trial #
Time (s) from lick onset
I
Pre-learning
Post-learning
Mean ΔF/F
Time (s) from lick onset
J
Mean ΔF/F dip
**
Pre-learning
Post-learning
K
ACh
DA
ΔF/F
Time (s) from lick onset
Latency (s)
***
DA dip
ACh late peak

<u>Figure 7:</u> Rapid dips in dopamine release in the anterior dorsal striatum to spontaneous lick onsets before learning and to Pavlovian 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 ∆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 ∆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 ΔF/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) ΔF/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. n.s., not significant (Kruskal-Wallis test, $p >0.05$). Each datapoint is the mean for one mouse. **H.** Left: Mean ∆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 ∆F/F for all bout onsets included in triggered average on left. **I.** Mean lick bout onset-triggered ∆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 ΔF/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) ∆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).

FIGURE 8


A
Fiber
AAV-iGluSnFR
Chat-cre mice
B
1mm
200 µm
C
iGluSnFR (470 nm)
Isosbestic control (405 nm)
ΔF/F
0.005
1s
D
Post-learning
Extinction 20%
Mean ΔF/F
Time (s) from cue onset
Light post-learning
Light extinction 20%
Trial #
ΔF/F
E
Post-learning
Extinction 20%
Mean ΔF/F
Normalized licking
Velocity (m/s)
Time (s) from cue onset

<u>Figure 8:</u> Cue-evoked glutamate release onto anterior striatum cholinergic interneurons does not change during extinction

A. Schematic showing strategy for single fiber optical measurements of bulk glutamate release selectively on anterior striatum CINs with iGluSnFR in ChAT-cre mice. **B.** Left: Fluorescence image of a coronal section of the anterior striatum showing iGluSnFR expression. Right: Zoomed image of the region in the white box on left showing restricted expression to cholinergic interneurons. **C.** Representative ∆F/F traces aligned to light cue onset for two trials in one mouse **D.** Top: Mean ∆F/F from one representative fiber from a single mouse aligned to light cue onset for trials post-learning (orange) and partial extinction (20% reward probability cue, green). Shaded regions, S.E.M. Bottom: Light cue-aligned ∆F/F for all trials included in the triggered averages at top. **E.** Mean ∆F/F (left), licking (middle), and velocity (right) (n = 8 sessions across 2 fibers in 2 mice) aligned to light (top) and tone (bottom) cue onsets for post-learning and 20% partial extinction phases. Shaded regions, S.E.M.

Methods

Mice

For all experiments adult male and female (n = 31 mice in total, postnatal 3-5 months, 24-30g), wild type (WT) C57BL6 /J (n=15, 10 males and 5 females, Jackson Labs, strain #00664) and heterozygous ChAT-cre mice (n=16, 8 males and 8 females, ChAT-IRES-Cre, Jackson Labs, strain# 006410) were used in this study. Mice were initially housed in groups and then individually housed following surgery, under standard laboratory conditions (20-26°C, 30-70% humidity; reverse 12-h/12 h light/dark cycle; light on at 9 p.m.) with *ad libitum* access to food and water, except during water scheduling. The number of mice used for each experiment and analysis is indicated in figure legends and main text. Experiments were conducted during the dark cycle. All animal care and experimental procedures were performed in accordance with protocols approved by the Boston University Institutional Animal Care and Use Committee (protocol no. 201800554) and in compliance with the Guide for Animal Care and Use of Laboratory Animals.

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

Array fabrication: The multifiber arrays to enable large-scale measurements of ACh (and DA) release across the 3D volume of the striatum were fabricated in-house as previously described[37] (**Fig. 1A**). Briefly, 51-99 optical fibers (50µm, 46µm core and 4µm cladding with 0.66NA, Fiber Optics Tech) were threaded into holes (55-60µm diameter) within a custom 3D-printed grid (3mm W x 5mm L, Boston Micro Fabrication) and secured with UV glue (Norland Optical Adhesive 61). Fibers were fixed at preset lengths from the bottom of the grid (measured under a dissection scope) and neighboring fibers were separated by at least 250 µm axially and 220 µm radially to target different locations throughout the striatum volume, ensuring no (or very minimal) overlap between fluorescence collection volumes [33–35]. Fiber ends on the non-implanted side were bundled and glued inside a polyimide tube (0.92mm ID, MicroLumen) and then cut with a sharp razor blade to a ~ 1cm length. The bundled fibers were then polished first on fine 6µm grained polishing paper followed by 3µm (Thorlabs) to create a uniform and smooth imaging surface and enable an efficient light transmission through each fiber (**Fig. 1A**).

Fiber localization: Following the array fabrication, a calibration procedure was performed to match individual fibers on the implanted side to their corresponding locations on the bundle surface as previously described[37]. At the end of neuronal recordings and behavioral experiments, mice were injected intraperitoneally with Euthasol (400 mg kg-1, Covertus Euthanasia III) then perfused intracardially with phosphate-buffered saline (PBS 1%, Fisher Scientific), followed by paraformaldehyde (4% in 1% PBS, Fisher Scientific). The mice were decapitated, and the ventral side of the skull removed to expose the ventral side of the brain. The intact implanted brains were post-fixed for 24h in 4% paraformaldehyde, rinsed with 1% PBS then transferred to a Lugol iodine solution (Carolina Scientific) diluted 1:3 in distilled water for 4 to 6 days. The diffusion of the Lugol solution to the brain enhances tissue contrast for computerized tomography scanning to enable fiber localization. The CT scanning and fiber localization were previously described[37]. Briefly, the 3-D CT scans of the intact implanted brains were registered to the The Allen Mouse Brain Common Coordinate Framework atlas via a semi-automated landmark-based approach[36,71]. The

fibers, bright in the CT scan, were then automatically identified via intensity thresholding, and the recording locations (ventral-most point) were mapped to their corresponding locations on the implanted grid and subsequently to their locations on the imaging surface. In addition to bringing all mice into a common coordinate space, registration with the atlas also enabled the recording locations to be automatically assigned atlas anatomical labels, which were further manually verified (**Fig. 1D**).

Stereotaxic viral injections and chronic multi-fiber array implants

Mice were anesthetized with isoflurane (3-4%) and placed in a stereotaxic frame (Kopf instruments) on an electric heating pad (Physitemp instruments) and administered buprenorphine extended release for pre-operative analgesia (3.25 mg kg-1 subcutaneous, Ethiqa XR). Following induction, isoflurane was held at 1-2% (in 0.8-1 L min^{-1} pure oxygen) and body temperature maintained at 37°C, throughout the surgical procedure. The scalp was shaved and cleaned with iodine solution prior to exposing the skull. For experiments to record extracellular acetylcholine release using multi fiber arrays, a large craniotomy was performed over the right hemisphere with a surgical drill (Midwest Tradition 790044, Avtec Dental RMWT) to expose the brain surface from -2.3 to 2 mm in the anterior-posterior (AP) direction and from 0.49 to 3.4 mm in the medial-lateral (ML) direction relative to bregma. AAV9-hSyn-ACh3.0[32] (WZ Biosciences), $2.07x10^{13}$ GC ml^{-1} diluted 1:2 in PBS was pressure-injected into the striatum of WT mice (n = 8) through a pulled glass pipette (tip diameter 30-50 µm) at 20-40 separate striatum locations chosen to maximize expression around fiber tips (200 nl at each location at a rate of 100 nl/min). For control experiments, a 1:2 mixture of AAV9-hysn-ACh3.0-mut[32] (WZ Biosciences), $2.54x10^{13}$ GC ml^{-1} was injected into the striatum of WT mice (n = 3) using the same strategy. For experiments to record extracellular dopamine release using multi-fiber photometry, AAV5-CAG-dlight1.3b[61] (Addgene, # 111067), 1.7x $x10^{13}$ GC ml^{-1} diluted 1:3 in PBS was injected into the striatum of WT mice (n = 4) at 10-40 total locations (200-800nl at each location) using the same procedure. Following injections, the multi-fiber array was mounted onto the stereotaxic manipulator, the dura gently removed, and the array slowly lowered into position. The craniotomy was sealed with a thin layer of Kwik-Sil (WPI), and the array was secured to the skull surface using Metabond (Parkell). To allow head fixation, a metal head plate and ring (Atlas Tool and Die Works) were next secured to the skull with Metabond, and the implant surface was covered with a mixture of Metabond and carbon powder (Sigma Aldrich) to reduce optical artifacts. The fiber bundle was protected by a cylindrical plastic tube, extending ~ 1-2 mm above the fiber bundle, and secured around the bundle using a mixture of Metabond and carbon powder.

To drive the suppression of ACh from cholinergic interneurons with tetanus toxin light chain (TeLC, **Fig. 6**), small, circular craniotomies were drilled bilaterally above the injection sites (from bregma, in mm; AP: 1, ML: ± 1.4). Then either pAAV2/8-hSyn-FLEX-TeLC-P2A-EYFP-WPRE[72] (Addgene, #135391), a gift from Bernando Sabatini's Lab, $5.14x10^{13}$ GC ml^{-1}) or ssAAV-5/2-hSyn1-dlox-TeTxLC_2A_NLS_dTomato(rev)-dlox-WRPE-hGHp (Viral Vector Facility University of Zurich, $4.1x10^{12}$ VG ml^{-1}) diluted 1:1 in PBS was bilaterally injected in the aDS of ChAT-cre mice (n = 6) at 4-12 sites per hemisphere (300nl/site at a rate of 100nl/min) at the following coordinates in mm; AP: 0.8, ML: ± 1.25, DV: -2.5 and -3; AP:1, ML: ±1.4, DV:-2.75 and -3. Control

ChAT-cre mice (n = 8 Chat-cre mice) were injected with saline using the same strategy. Following the injections, the craniotomies were sealed with Kwik-Sil (WPI), and the skull was sealed with Metabond (Parkell) and a metal head plate. To measure extracellular glutamate release into cholinergic interneurons (**Fig. 8**), craniotomies were drilled above the injection sites in the right hemisphere (from bregma, in mm; AP: 1, ML: 1.4). A 1:1 mixture in PBS of the genetically encoded glutamate sensor AAV9.hSyn-FLEX.8F-iGluSnFR.A184S[62] (Adgene, #106174), $1.8x10^{13}$ GC ml^{-1} was injected in aDS of ChAT cre mice (n = 2) at 6 sites (300nl/site at a rate of 100nl/min) at the following 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, -3.25 and -3.75. Then, a 100 µm core diameter optical fiber (MFC_100/125- 0.37NA) attached to a zirconia ferrule (Doric) was slowly lowered into the aDMS (AP:1, ML:1.4) to a final depth of 3 mm from bregma. The craniotomies were sealed with Kwik-Sil (WPI), the optical fiber and a head plate were secured to the skull with Metabond (Parkell). After the surgeries, mice were placed in a cage with a heating pad and received postoperative injections of meloxicam (5 mg kg^{-1} subcutaneous, Covertus) and 1 mL of saline per day subcutaneously for 4 days after surgery. Mice were individually housed and allowed to recover in their cages for at least 2 weeks after surgery.

Pavlovian conditioning task and behavior setup

Behavioral apparatus: One week before starting the Pavlovian conditioning training and photometry recordings, mice were placed on a water schedule, receiving 1 ml of water per day, and were maintained at 80-85% of their initial body weight for the duration of the experiments. Three to four days prior to training, mice were habituated to head fixation on the spherical styrofoam treadmill[40] (Smoothfoam, 8in diameter, **Fig. 1B**). The behavioral setup has been described in detail previously[37]. Briefly, mice were free to locomote on the treadmill during all experiments and the ball rotation in pitch, yaw, and roll directions was measured using optical computer mice (Logitech G203) through an acquisition board (NIDAQ, PCIe 6343). Water rewards (5µL/reward) were dispensed through a water spout operated by an electronically controlled solenoid valve (#161T012, Neptune Research), mounted on a post a few mm away from the mice's mouths. Tongue spout contacts (a proxy for licking) were monitored by a capacitive touch circuit connected to the spout and confirmed with live video taken from a camera positioned to capture orofacial movements (Blackfly S USB3, BFS-U3-16S2M-CS, Teledyne Flir).

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

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

Multi-fiber photometry recordings

Fluorescence measurements from the multi-fiber arrays were conducted using a custom-built microscope (**Fig. 1B**) mounted on a 4' W x 8' L x 12'' thick vibration isolation table (Newport). Details of the microscope were described previously[37]. Excitation light for the fluorescent sensors (ACh3.0, ACh4.3 mut, iGluSnFR and dLight 1.3b) was provided by two high power LEDs (470nm and 405nm; Thor labs, No. SOLIS-470C, SOLIS-405C). Excitation light was bandpass filtered (Chroma No. ET405/10 and ET473/24) then coupled into a liquid light guide (Newport No. 77632) with lenses (f = 60mm and 30mm, Thor labs No. LA1401-A and LA1805) and a collimating beam probe (Newport No. 76600). The liquid light guide was connected to a filter cube on the microscope, directing excitation light into the back aperture of the microscope objective (10x, 0.3NA, Olympus Model UPLFLN10X2) via a dichroic beam splitter (Chroma Model 59009bs). The light power at the focal plane of the objective was adjusted to be within the range of 80-85mW, resulting in a power of 1.6~2mW/mm^2 at the fiber tips[37]. Emission light was bandpass filtered (Chroma, No 525/50m) and focused with a tube lens (Thor labs, No TTL165-A) onto the CMOS sensor of the camera (Hamamatsu, Orca Fusion BT Gen III), creating an image of the fiber bundle (**Fig. 1B**). To enable precise manual focusing, the microscope was connected to a micromanipulator (Newport Model 96067-XYZ-R) and mounted on a rotatable arm extending over the head-fixation setup to facilitate positioning of the objective above the imaging surface over the mouse head. Imaging data acquisition was performed using HCImage live (HCImage live, Hamamatsu). Single wavelength excitation was carried out with continuous imaging at 30Hz (33.33ms exposure time), via internal triggering. Dual wavelength excitation was performed in a quasi-simultaneous externally triggered imaging mode, where the two LEDs were alternated and synchronized with imaging acquisition via 5V digital TTL pulses. 470 nm excitation was alternated with 405 nm excitation at either 36Hz (20ms exposure time) or 22Hz (33.33ms exposure time) to achieve a frame rate of 18 Hz or 11 Hz for each excitation wavelength, respectively. Recordings acquired at different sampling rates were downsampled or upsampled using a 1-D interpolation with Matlab's *interp1* function using the spline method. A custom MATLAB software controlled the timing and duration of TTL pulses through a programmable digital acquisition card (NIDAQ, National Instruments PCIe 6343). Voltage pulses were transmitted to the NIDAQ from the camera following the exposure of each frame to confirm proper camera triggering and to synchronize imaging data with behavior data.

Statistics and reproducibility

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

Multi-fiber photometry signal preprocessing

The acquired time series videos of the fiber bundles were first motion-corrected using a whole-frame cross-correlation algorithm described previously[40,73] then visually inspected to confirm post-correction image stability. Circular regions of interest (ROI, ~ 25μm diameter) were manually selected for each fiber. The resulting set of ROIs comprised a mouse-specific ROI template, which was then fit and applied to each subsequent imaging video, enabling the identity of each ROI to remain consistent across multiple recording sessions. To determine the change in fluorescence ΔF/F, the mean fluorescence extracted from each ROI was normalized to a baseline, which was defined as the 8th percentile fluorescence over a 30-s sliding window[37]. To remove low frequency artifacts, the ΔF/F signals were high-pass filtered using a finite impulse response filter with a passband frequency set at 0.3Hz. This frequency was determined based on the observed differences in the dynamics of the ACh signal compared to the control signals (ACh-mut and the isosbestic 405 nm LED signal, **Supplementary Fig. 1**). Most analyses were conducted on non z-scored or peak normalized ΔF/F values in order to identify relative differences in signal magnitude across task and training phases. Changes in overall signal magnitude over training due to changes in sensor expression or fiber collection efficiency were accounted for by examining the stability of signals to unpredicted water reward for each fiber (**Supplementary Fig. 1**).

Quantification and statistical analysis

Definition of learning phases

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

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

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

indices of both cues were significantly higher than the first session were considered post-learning sessions (**Fig. 2D**, **Supplementary Fig. 3A**). Sessions where only one cue had a significantly elevated lick index were omitted from analysis. For the extinction phase (**Fig. 4**), sessions were included in analyses if the lick indices of the 20% probability cue were significantly (Wilcoxon rank-sum test, $p<0.05$) lower than lick indices for the same cue on the last day of the post-learning phase or the preceding 80% probability session (for reversal sessions). The fraction of time spent licking during the cue period (**Fig. 2E**) was calculated across the entire 3-s window after cue onset, and the fraction of time spent licking during the ITI period was calculated across the entire ITI period, excluding 0.5s before cue onset, 3s after cue offset, and 6s after any unpredicted reward delivery. Spontaneous lick bout onsets in the ITI (**Fig. 5**) were defined as the first lick of lick bouts that are not preceded by any licking for at least two seconds. For lick analyses in TelC experiments (**Fig. 6**), all the extinction sessions (20% reward probability for both light and tone cues combined) were compared with sessions of pre- and post-learning (Two-way ANOVA followed by Tukey *post hoc* analysis to account for multiple comparisons within and between TelC and control mice, $p<0.05$ was considered statistically significant).

Relationships of signal with velocity and acceleration

Analog signals from the optical mice were converted to m/s[37], and the pitch and roll were combined to compute a total velocity calculated as: $\sqrt{pitch^2 + roll^2}$. The velocity traces were then smoothed using a Savitzky-Golay moving average filter with a moving window of 250 ms and a 2nd degree polynomial. Acceleration traces were derived from the smoothed velocity (acceleration = Δvelocity/Δtime), and further filtered using the same Savitzky-Golay moving average filter parameters.

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

Equation 1:

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

GLM training, where $signal_{\Delta F/F_{filtered}}$ is the filtered ΔF/F of the ITI periods, v_1/v_2 and a_1/a_2 are positive/negative phase-shifted velocities and accelerations respectively, β_{v1}/β_{v2} are the correlation coefficients of positive/negative phase-shifted velocities, β_{a1}/β_{a2} are the correlation coefficients of positive/negative phase-shifted accelerations, ε is the error term.

Equation 2:

$$signal_{velocity} = \widehat{\beta_0} + \widehat{\beta_{v1}} \times v_1 + \widehat{\beta_{v2}} \times v_2 + \widehat{\beta_{a1}} \times a_1 + \widehat{\beta_{a2}} \times a_2$$

$signal_{velocity}$is the estimated contribution of velocity and acceleration to the ΔF/F signal, v_1/v_2 and a_1/a_2 are the same phase-shifted velocities and accelerations as described above, $\hat{\beta}$ are the estimated correlation coefficients generated by the GLM. Non-significant $\hat{\beta}$, i.e. with a p > 0.05, were set to 0 and were not included in the estimated corresponding velocity or acceleration contribution.
The final ΔF/F was computed by subtracting the velocity contribution ($signal_{velocity}$) from the filtered signal ($\Delta F/F_{filtered}$).

To further address the possibility that variations in treadmill velocity at cues or rewards contribute to ΔF/F signal changes, signals and velocity/acceleration were compared for different learning phases for each fiber (**Fig. 4M**, **Fig. 7G**, **Supplementary Fig. 4A-G**, **Supplementary Fig. 5A-G**, **Supplementary Fig. 6G**, **Supplementary Fig. 9B-J**). Periods of continuous deceleration were first defined in peri-event windows as consecutive bins with negative acceleration values. Peak deceleration was defined as the minimum value of acceleration during any single continuous deceleration period, in which the total velocity change exceeded a threshold of 8 cm/s. Total velocity changes following cue or reward were defined as the maximum change of velocity during the peri-event windows. Total velocity change and peak decelerations were calculated from cue onset to 0.6s after, and from -0.2 to 0.6s after reward consumption. For non-contingent spontaneous licks during ITI periods, total velocity change and peak decelerations were calculated from -0.5 to 0.1s after lick onset. Large and small deceleration trials were defined as trials where the peak deceleration was greater than the 80th percentile or less than the 20th percentile of the total peak deceleration distribution across all mice, respectively (**Supplementary Fig. 4A-G**, **Supplementary Fig. 5A-G**, **Supplementary Fig. 9B-J**).

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

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

Significant positive or negative changes in ACh release were first identified around each event (cue onset, reward delivery, lick bout onset) within a peri-event window (0-1.5s for cue, -0.5 -1.5s for reward, and -0.5-1s for lick onsets). Significance was determined as a signal change exceeding 3 standard deviations above (peaks) or below (dips) the mean of the ITI signal in a 1s window prior to the event window. Local maximum (peaks) and minimum (dips) ΔF/F points were defined for each fiber and the timing of these significant points was plotted in histograms (**Fig. 1G**, **Fig. 2G**, **Fig. 3B**, **Fig. 4D**, **Fig. 5D**, **Supplementary Fig. 3C**, **Supplementary Fig. 6A**). Based on the clear multi-phasic timing distributions from the histograms, windows were defined for each event to define significant early peak, dip, and late peak components. The windows were as follows for early peak, dip, and late peak respectively: cue onset: 55ms to 444ms, 333ms to 833ms, 500ms to 1222ms; reward consumption: -200ms to 400ms, 200ms to 800ms, 500ms to 1200ms; ITI lick bout onset: -270ms to 230ms, 0ms to 600ms, 200ms to 800ms. For maps

showing the presence of each set of components for each fiber, additional criteria were included that a signal must exceed 3 standard deviations relative to the ITI for at least 2 consecutive 0.056ms time bins and reach a peak or trough of at least 0.005 ΔF/F or -0.005 ΔF/F respectively to isolate only the largest signal changes.

Differences in the mean event-evoked ΔF/F between different phases of learning for the three ACh signal components were calculated for each fiber by comparing the peak and dip values across single trials within each phase using a non-parametric unpaired Wilcoxon rank sum test. P-values <0.01 were considered statistically significant (**Fig. 2I**, **Fig. 3D**, **Fig. 4J**, **Supplementary Fig. 3F**, **Supplementary Fig. 4H**, **Supplementary Fig. 5H**, **Supplementary Fig. 6D**, **Supplementary Fig. 7C** and **D**, **Supplementary Fig. 8A**), except for the difference in mean ITI lick bout-evoked ΔF/F, where p-values <0.025 were considered significant (**Fig. 5F**, **Supplementary Fig. 9A** and **K**).

Relative signal timing calculations

Latencies for ACh peaks or dips were determined as the time between the event onset and the time of the minimum or maximum signal amplitude in the triggered average for a given signal component (early/late peak, dip). To determine the relative timing between the increase in light and tone cue-evoked late peak ΔF/F and the subsequent decrease in lick index following partial extinction (**Fig. 4K,L**), lick indices and late peak amplitude values for each fiber across trials were smoothed using Matlab's *smooth* function with the lowess method (local regression using weighted linear least squares and a first-degree polynomial model with a moving window of 30 trials). Next, Matlab's *cusum* function with a climit of 5 and mshift of 2 was used to identify the change point for lick indices and late peak ΔF/F (a decrease in lick index and an increase in late peak). The *cusum* algorithm detects small incremental changes by maintaining cumulative sums of detectable positive or negative shifts from the mean of each data point in the sequence. The threshold for detectable shifts was determined by the product of mshift and the standard deviation of the sequence. Significance was established when the upper or lower cumulative sum of shifts exceeded a threshold based on the product of climit and the standard deviation of the sequence (**Fig. 4K** and **I**, **Supplementary Fig. 6E** and **F**).

Dopamine changes during learning and extinction

Measurements of DA release (dLight1.3b ΔF/F, **Fig. 7**) were performed and analyzed similarly to ACh data. For the analysis of DA signals, 29 fibers in the aDS were selected from a total of 150 fibers within the striatum (out of 324 implanted fibers), based on the criteria AP>0mm, ML<2mm and DA<-4mm. These fibers were selected from 4 mice (10, 6, 6, 7 fibers in each mouse respectively) implanted with arrays. Other fibers were ignored for this study and anatomical selection criteria was defined only with respect to the regional ACh signaling patterns in the ACh3.0 group and blindly with respect to the DA signals. Dips in DA release were identified as the minimum values of the trial averaged dLight1.3b ΔF/F within a time window 0.2 to 1.5s after cue onset or spontaneous unrewarded lick bouts. The latencies of DA dips to spontaneous unrewarded licks bouts in the ITI were calculated as the time elapsed between the ITI lick bout onsets and the minimum ΔF/F, averaged across trials (**Fig. 7I**). The latencies between the

increase in cue-evoked DA dips and decrease in lick indices during extinction was determined using the same analysis as for ACh (see above, **Fig. 7E**).

Histology

At the end of glutamate recordings and TelC behavior experiments, mice were injected intraperitoneally with Euthasol (400 $mgkg^{-1}$, Covertus Euthanasia III) then perfused intracardially with phosphate-buffered saline (PBS 1%, Fisher Scientific), followed by paraformaldehyde (PFA 4% in 1% PBS, Fisher Scientific). The brains were post-fixed overnight in 4% PFA dissolved in PBS and then transferred to a solution of 40% sucrose in 1% PBS (until the brains sank). The brains were sliced (50µm thickness) with a cryostat (Leica CM3050 S). The coronal sections were then mounted on super frost slides and cover slipped with Vectashield antifade mounting medium (Vector Laboratories, H-1900). TelC (**Fig. 6B**) and iGluSnFr (**Fig. 8B**) fluorescence were not immuno-enhanced; confocal images were acquired on a Zeiss LSM 800 laser scanning confocal microscope.

Acknowledgements

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

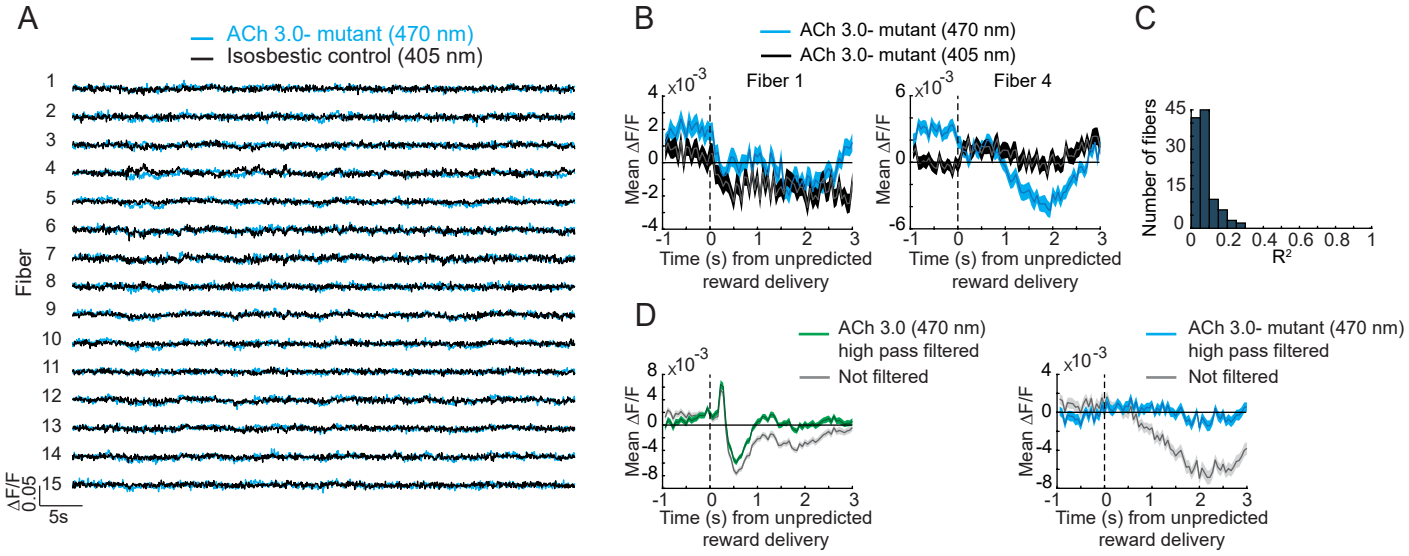
Author contributions

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

Competing Interests: The authors declare no competing interests

Supplementary Figures

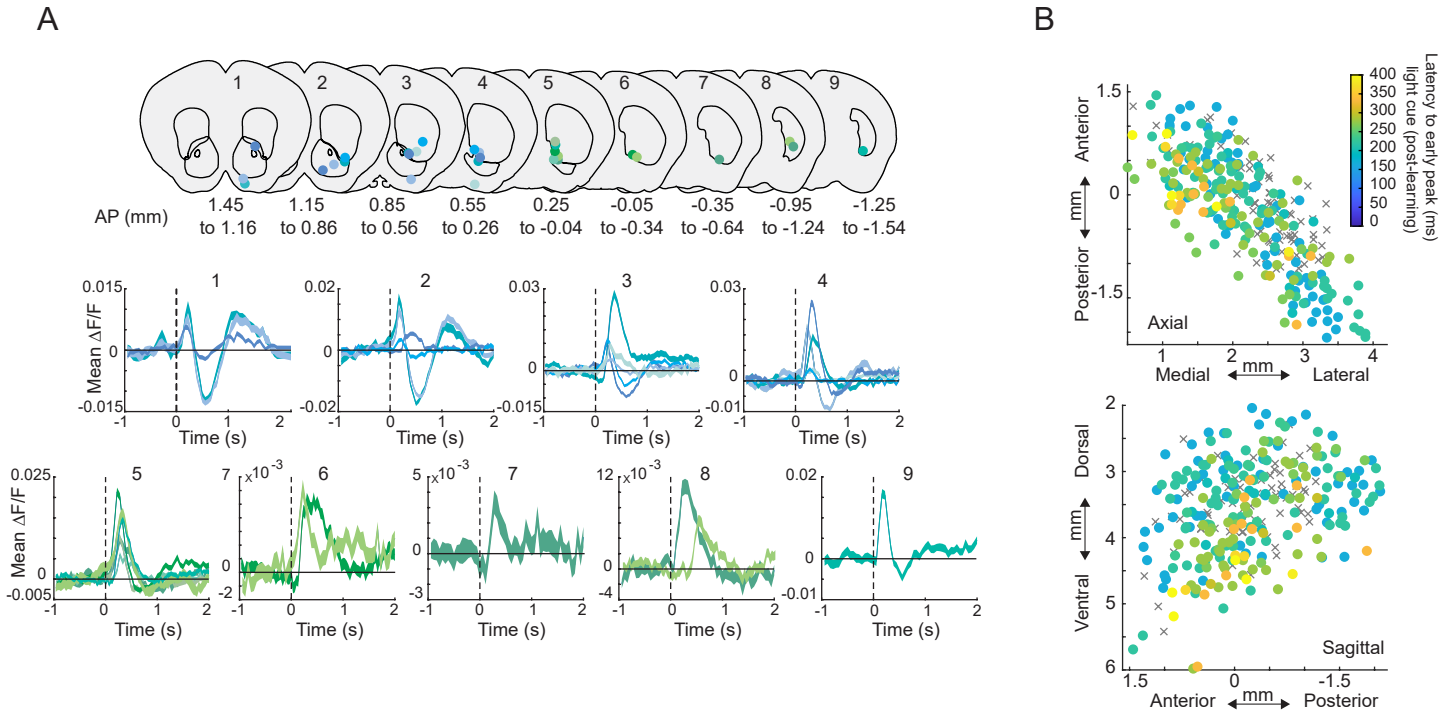
SUPPLEMENTAL FIGURE 1
Related to Figure 1



<u>Supplementary Figure 1:</u> Controls for hemodynamic and motion artifacts

A. Fluorescence traces ($\Delta F/F$) from 15 representative fibers in a mouse expressing a non-functional mutated version of the ACh3.0 sensor with alternating 405 and 470nm illumination. **B.** Mean $\Delta F/F$ aligned to unpredicted reward delivery from two fibers in the mouse shown in (A). Note that the small sustained decrease in fluorescence detected with 470nm illumination is only clearly detected with 405nm illumination in the right fiber. Shaded region, S.E.M. **C.** Histogram of Pearson's correlation R^2 between $\Delta F/F$ fluorescence traces obtained with quasi-simultaneous 405nm and 470nm illumination in mutant ACh sensor expressing mice (n = 3). **D.** Mean unpredicted reward triggered $\Delta F/F$ from a single fiber in functional ACh3.0 sensor (left) and mutant ACh sensor (right) expressing mice before and after 0.3Hz high pass filtering. Shaded region, S.E.M. Note that the small, slow artifactual decrease in the mutant sensor recording is largely eliminated with filtering, but the rapid reward triggered release measured with the functional sensor is preserved.

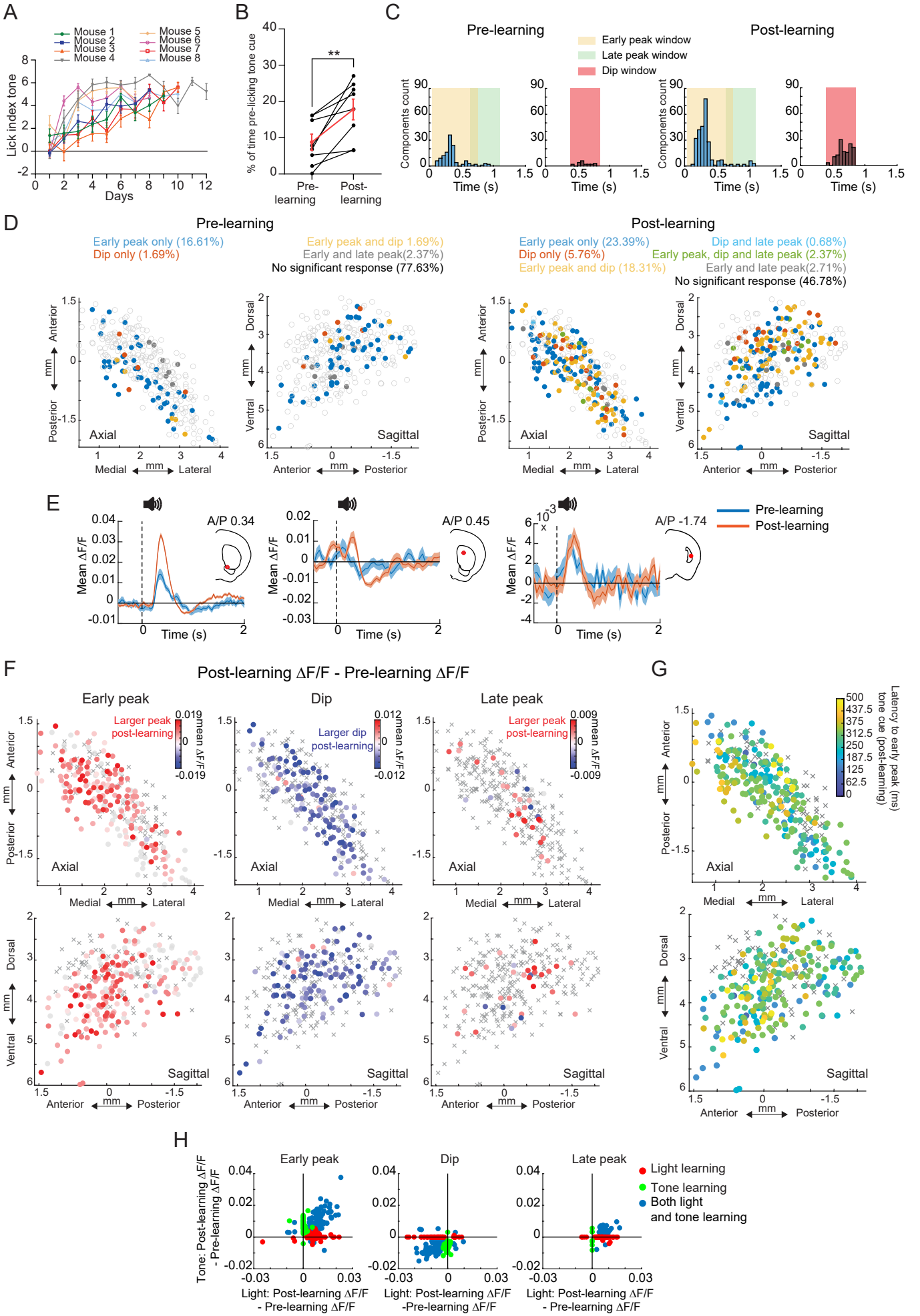
SUPPLEMENTAL FIGURE 2
Related to Figure 2



<u>Supplementary Figure 2:</u> Micro-organization of cue-evoked ACh release to conditioned light cues along the ventral anterior-posterior axis

A. Light cue onset-triggered $\Delta F/F$ averages (bottom) for 29 ventral striatum fibers in 8 mice on post-learning trials (n = 40-100 trials). Colors of the traces correspond to the fiber locations shown in the coronal plane schematics on top, where blue colors correspond to more anterior locations and green colors to more posterior locations. Shaded regions, S.E.M. **B.** Map of latencies to early peaks to light cue onsets post-learning for all fibers. Xs are fibers with no significant early peak.

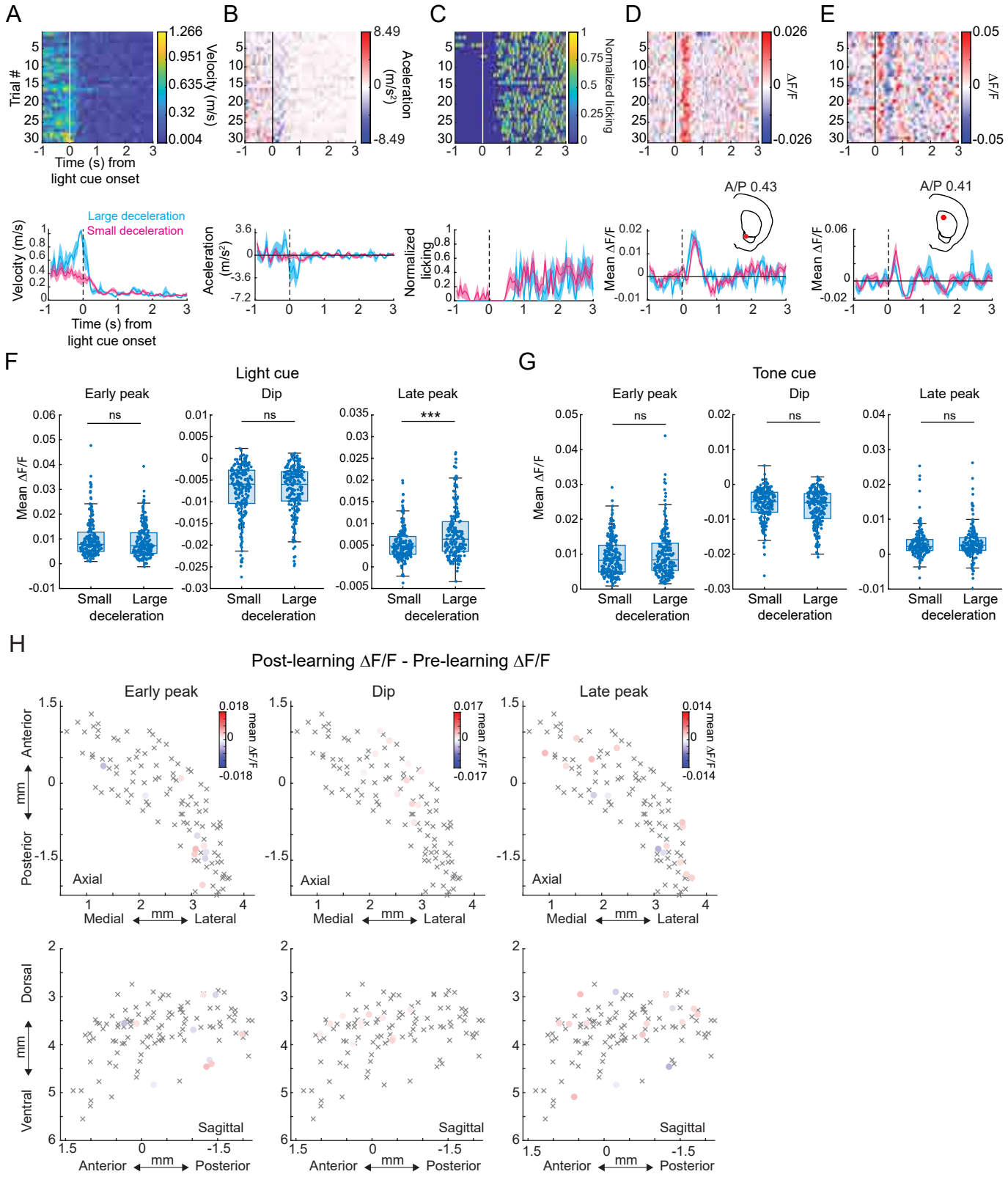
SUPPLEMENTAL FIGURE 3
Related to Figure 2



<u>Supplementary Figure 3:</u> Region specific, bi-directional changes in multi-phasic ACh release to conditioned tone cues during Pavlovian learning

A. Mean lick index on tone cue trials for each day after the start of Pavlovian training for each mouse. **B.** Mean percent time spent licking during the 3s tone cue interval prior to reward for each individual mouse for sessions pre- and post-learning (as indicated by the overall lick index, see Methods). Red points are means across all mice (n = 8), significant difference between pre- and post-learning assessed by a two-tailed non-parametric Wilcoxon matched-pairs signed rank test (**$p<0.01$). **C.** Histogram of the number of fibers expressing a significant early peak, dip, and late peak response component to tone cue onset pre (left) and post (right) learning ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods). **D.** Maps indicating the presence of each combination of signal components to light cue onset for each fiber (circle) pre and post learning. Empty circles are fibers with no significant response. **E.** Mean $\Delta F/F$ for 3 representative fibers aligned to the tone cue onset for trials pre (blue) and post (orange) learning. Shaded regions, S.E.M. Red dots in insets indicate the fiber locations in the coronal plane. Examples on the left and center are the same fibers shown in Figure 2F. **F.** Maps showing the difference in mean tone-cue evoked $\Delta F/F$ with learning (post – pre) for the three signal components. Gray circles, no significant difference; Xs, no significant component. The significance of each component for each fiber was calculated using the Wilcoxon rank-sum test, two-tailed ($p < 0.01$). **G.** Map of latencies to early peaks to tone cue onsets post-learning for all fibers. Xs are fibers with no significant early peak. **H.** Difference in mean tone cue-evoked $\Delta F/F$ with learning (post – pre) vs the difference in mean light cue-evoked $\Delta F/F$ with learning of the three components. Each dot is one fiber and color indicates whether the $\Delta F/F$ difference with learning is significant ($p<0.01$, Wilcoxon rank-sum test, two-tailed) for the tone, light, or both.

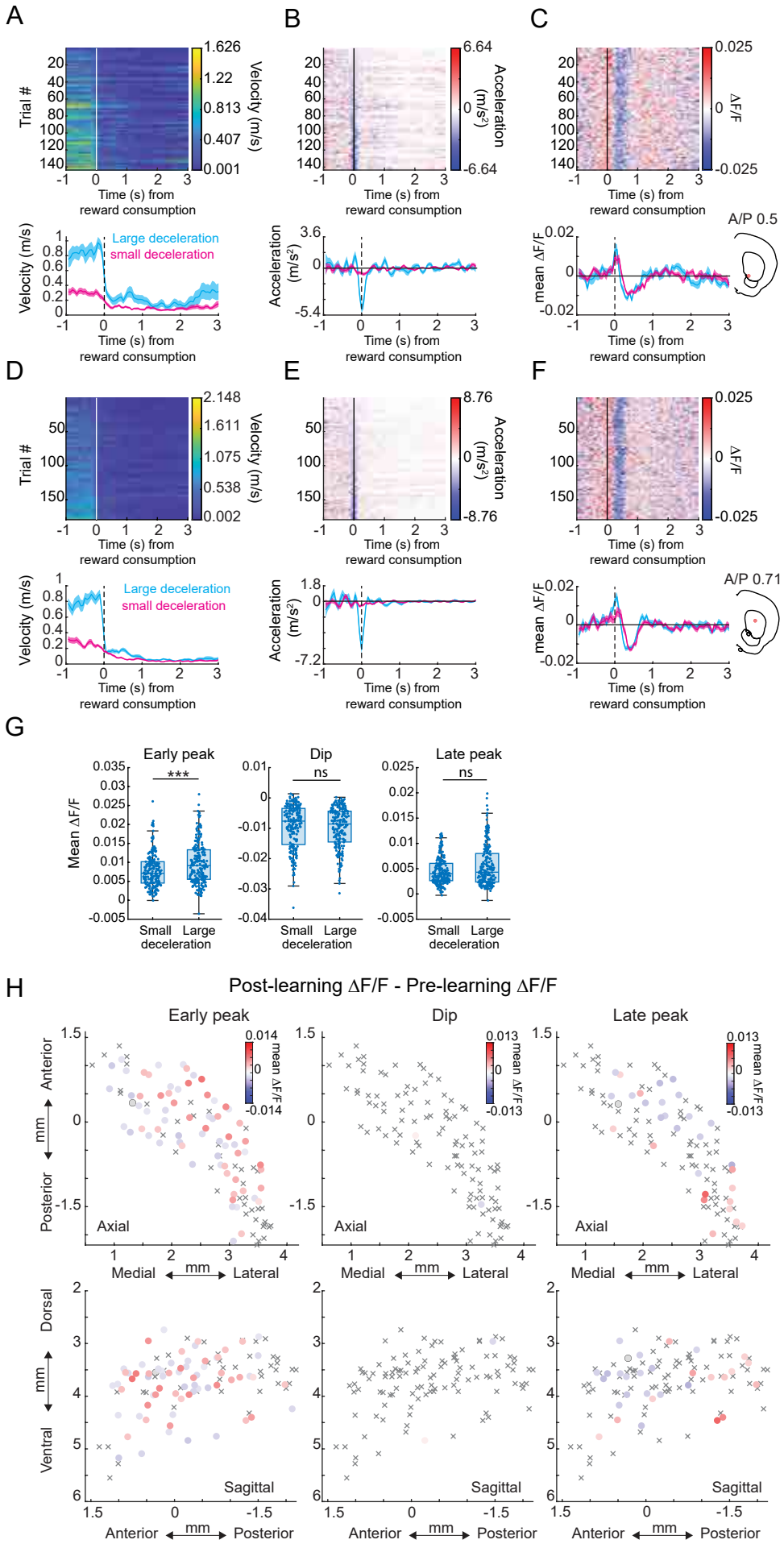
SUPPLEMENTAL FIGURE 4
Related to Figure 2



<u>Supplementary Figure 4:</u> Learning related changes in striatum-wide cue related ACh release cannot be explained by locomotion changes or motion/hemodynamic artifacts

A. Top: Total treadmill velocity aligned to light cue onsets for all trials in a single example session post-learning sorted by deceleration magnitude following cue onset. Bottom: Mean cue triggered average of treadmill velocity for trials with large decelerations (>20 percentile of decelerations from the distribution of decelerations across all mice, n = 4 trials, blue) and small decelerations (<20 percentile, n = 8 trials, magenta). Shaded region, S.E.M. **B.** Same as (A) for acceleration. **C.** Same as (A) for spout licking. **D.** Top: Light cue-aligned $\Delta F/F$ for a single fiber (location indicated in the coronal section inset) across all trials for the same session as A-C sorted by deceleration magnitude as in (A). Note that the magnitude of the cue-evoked $\Delta F/F$ does not vary with deceleration magnitude. Bottom: Mean cue triggered average $\Delta F/F$ for trials with large and small decelerations. Shaded region, S.E.M. **E.** Same as (D) for another example fiber. **F.** Box and whisker plots showing the mean light cue-evoked $\Delta F/F$ for each component (early peak, dip and late peak) for trials with small and large decelerations across all fibers (blue dots). Box bounds indicate the 25th and 75th percentiles and the whiskers represent the minima and maxima. Significance calculated between large and small decelerations for each component across fibers using a two-tailed Wilcoxon rank-sum test. ***: $p < 0.001$. n.s.: not significant. **G.** Same as (F) for tone cue. **H.** Maps showing the difference in mean light-cue evoked $\Delta F/F$ with learning (post – pre) for the three signal components in mice expressing non-functional mutant ACh sensor (n = 3 mice). Each circle is a single fiber. Gray circles, no significant difference; Xs, no significant component. Significance of each component for each fiber was assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$).

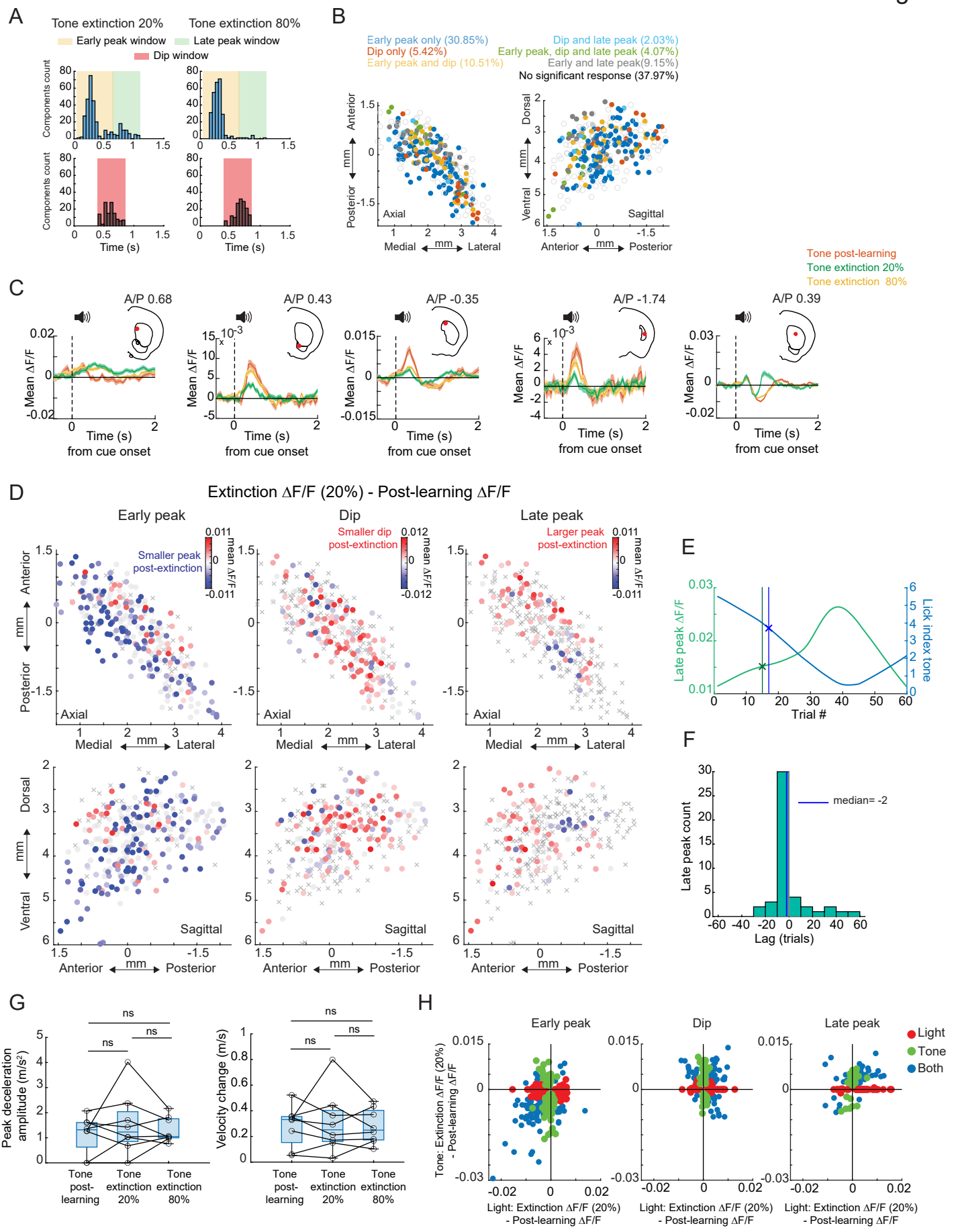
SUPPLEMENTAL FIGURE 5
Related to Figure 3



<u>Supplementary Figure 5:</u> Learning related changes in striatum-wide reward related ACh release cannot be explained by locomotion changes or motion/hemodynamic artifacts

A. Top: Total treadmill velocity aligned to reward consumption for all trials in a single example session post-learning sorted by deceleration magnitude following reward consumption. Bottom: Mean reward triggered average of treadmill velocity for trials with large decelerations (>20 percentile of decelerations from the distribution of decelerations across all mice, n = 13 trials, blue) and small decelerations (<20 percentile, n = 23 trials, magenta). Shaded region, S.E.M. **B.** Same as (A) for acceleration. **C.** Top: Reward aligned $\Delta F/F$ for a single fiber (location indicated in the coronal section inset) across all trials for the same session as A-B sorted by deceleration magnitude as in (A). Note that the magnitude of the reward evoked $\Delta F/F$ does not vary with deceleration magnitude. Bottom: Mean reward triggered average $\Delta F/F$ for trials with large and small decelerations. Shaded region, S.E.M. **D-F.** Same as A-C for a second representative fiber in a different mouse with large decelerations (>20 percentile of decelerations, n = 24 trials, blue) and small decelerations (<20 percentile of decelerations, n = 31 trials, magenta). **G.** Box and whisker plots showing the mean reward evoked $\Delta F/F$ for each component (early peak, dip and late peak) for trials with small and large decelerations across all fibers (blue dots). Box bounds indicate the 25th and 75th percentiles and the whiskers represent the minima and maxima. Significance calculated between large and small decelerations for each component across fibers using a two-tailed Wilcoxon rank-sum test. ***: $p < 0.001$. n.s.: not significant. **H.** Maps showing the difference in mean reward-evoked $\Delta F/F$ with learning (post – pre) for the three signal components in mice expressing non-functional mutant ACh sensor (n = 3 mice). Each circle is a single fiber. Gray circles, no significant difference; Xs, no significant component. Significance of each component for each fiber was assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$).

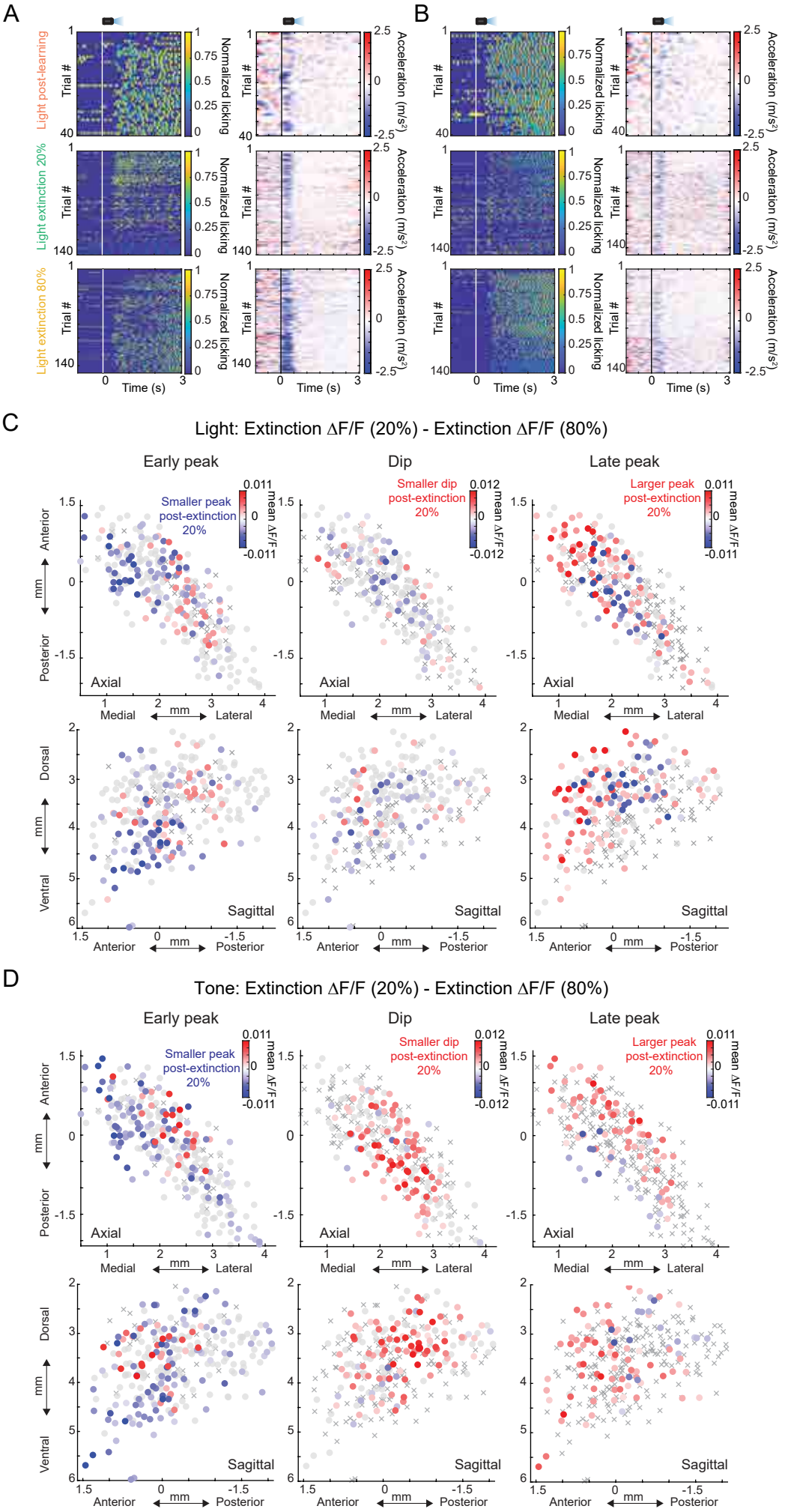
SUPPLEMENTAL FIGURE 6
Related to Figure 4



<u>Supplementary Figure 6:</u> Tone cue-evoked changes in striatal ACh release during cue-specific extinction

A. Histogram of the number of fibers expressing a significant ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods) early peak, dip, and late peak response component to tone cue onset associated with 80% or 20% reward probability. **B.** Maps (axial, left; sagittal, right) indicating the presence of each combination of signal components to 20% reward probability tone cue onset for each fiber (circle). Empty circles are fibers with no significant response. **C.** Mean $\Delta F/F$ for 5 representative fibers aligned to tone cue onset for trials where tone was associated with 100% (post-learning), 80%, and 20% (extinction) probabilities. Shaded regions, S.E.M. across trials. Red dots in the insets indicate the fiber locations in the coronal plane. **D.** Maps showing the difference in mean tone cue-evoked $\Delta F/F$ between the post-learning (100%) and 20% reward probability phases for the three signal components. Each circle is one fiber. Gray circles, no significant difference; Xs, no significant component. Significance differences for each component for each fiber were assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$). **E.** Mean tone 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. **F.** Histogram showing the relative latencies across fibers between the significant increase in tone 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. **G.** Box and whisker plots showing the peak decelerations and velocity changes across all mice (n=8 mice) following tone cue onsets. n.s., not significant (Kruskal-Wallis test, $p > 0.05$). Each datapoint is the mean for one mouse. **H.** Differences in mean cue evoked $\Delta F/F$ between 20% and 100% phases for light vs tone for each signal component. Each dot is one fiber and the color indicates whether the $\Delta F/F$ difference is significant ($p<0.05$, Wilcoxon rank-sum test, two-tailed) for the tone, light, or both.

SUPPLEMENTAL FIGURE 7
Related to Figure 4



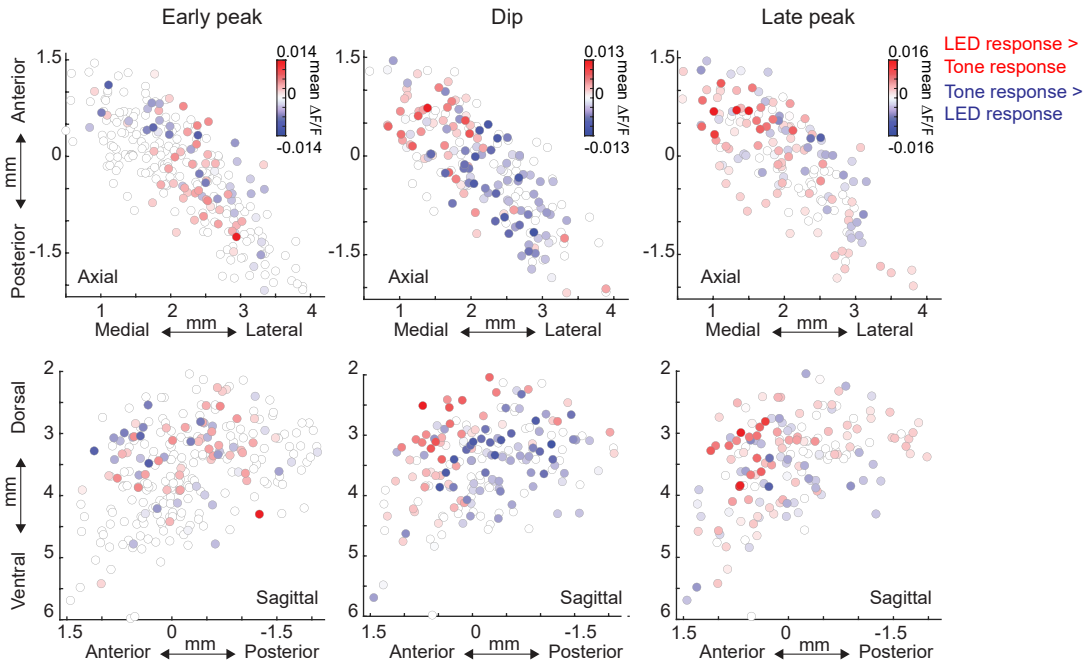
<u>Supplementary Figure 7:</u> Trial by trial behavior and comparison of ACh release for high and low probability cues during partial extinction sessions

A. Light cue onset-aligned spout licking (left) and treadmill acceleration (right) for all trials across the post-learning, 20% reward probability, and 80% reward probability phases (in this order) corresponding to the same trials shown in Figure 4F-G. **B.** Same as (A) corresponding to the trials shown in Figure 4H-I. **C.** Maps showing the difference in mean light cue-evoked $\Delta F/F$ between the 80% and 20% reward probability phases for the three signal components. Each circle is one fiber. Gray circles, no significant difference; Xs, no significant component. Significance differences for each component for each fiber was assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$). **D.** Maps as in (C) for tone cues.

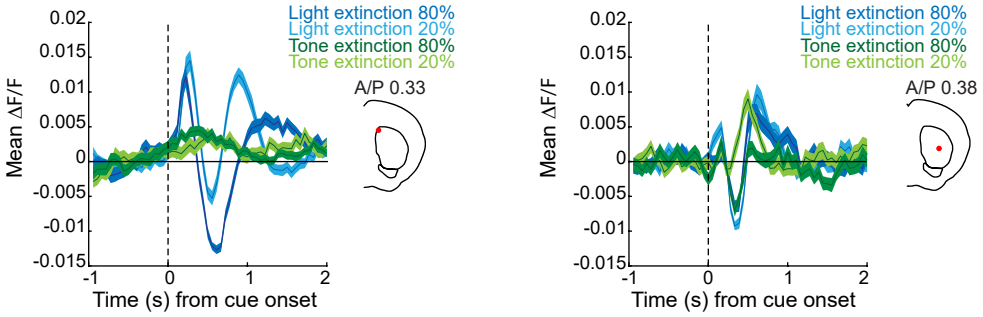
SUPPLEMENTAL FIGURE 8
Related to Figure 4

A

Light (Extinction ΔF/F (20%) - Post-learning ΔF/F) - Tone (Extinction ΔF/F (20%) - Post-learning ΔF/F)



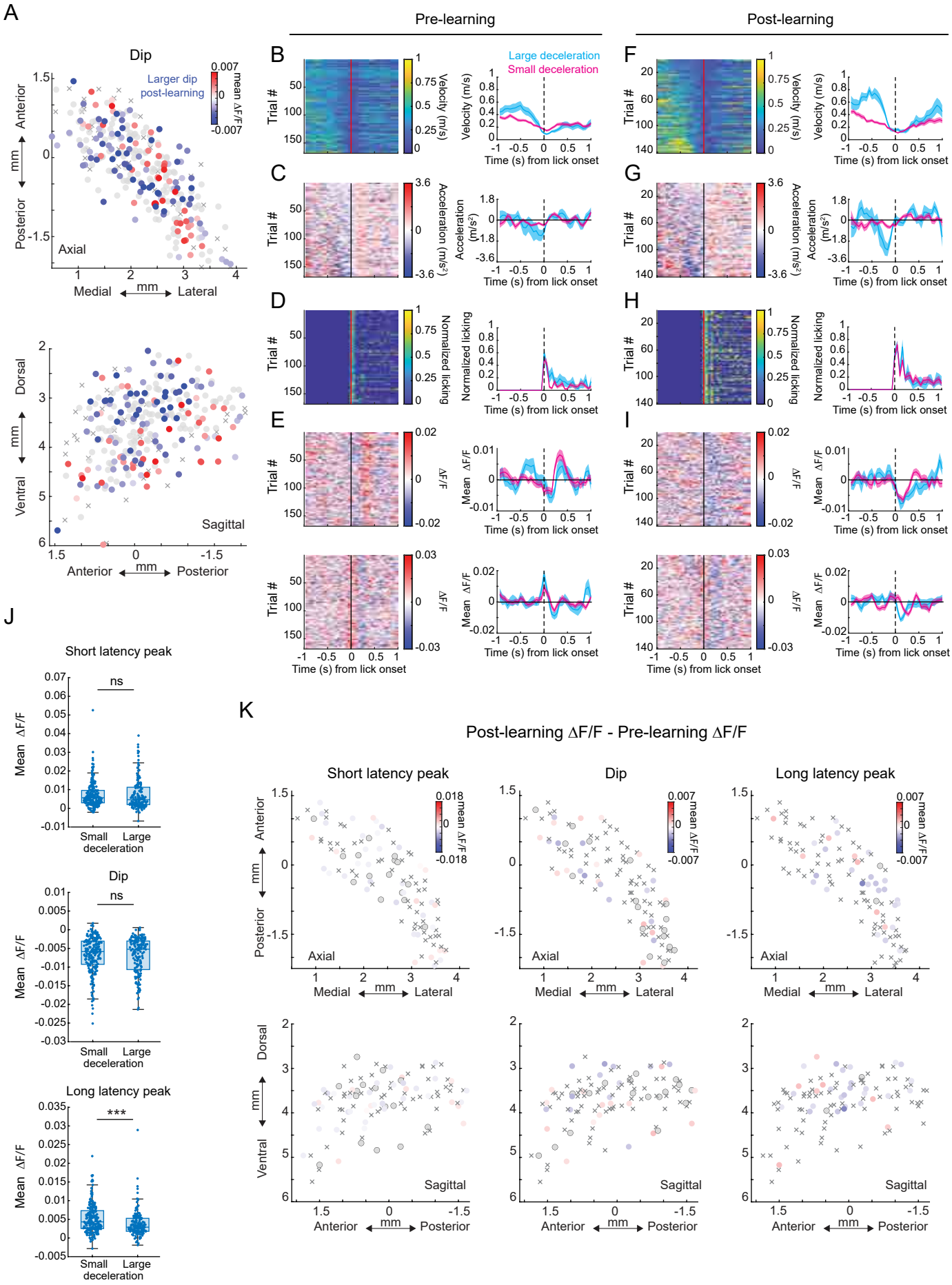
B



<u>Supplementary Figure 8:</u> Comparison between light and tone evoked ACh release during partial extinction

A. Maps showing differences between mean light and tone cue-evoked $\Delta F/F$ to the 20% probability cue relative to post-learning for each signal component. Each circle is one fiber. Red colors indicate stronger response for the light, blue for tone. White circles, no significant difference. Significance differences for each component for each fiber were assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$). **B.** Mean $\Delta F/F$ for two representative fibers aligned to tone or light cue onset for trials where each cue was associated with 20% or 80% reward probability during the partial extinction phase. Shaded regions, S.E.M. across trials. Red dots in the insets indicate the fiber locations in the coronal plane.

SUPPLEMENTAL FIGURE 9
Related to Figure 5



<u>Supplementary Figure 9:</u> Learning related changes in unrewarded, spontaneous lick-evoked ACh release cannot be explained by locomotion changes or motion/hemodynamic artifacts

A. Map showing the difference in mean ITI lick bout-evoked ∆F/F between pre and post learning for the dip as in Fig. 5F. **B.** Left: Total treadmill velocity sorted by deceleration magnitude aligned to spontaneous unrewarded ITI lick bout onsets for all onsets in a single session pre-learning. Right: Mean ITI lick bout onsets triggered average of treadmill velocity for trials with large decelerations (>20 percentile of decelerations from the distribution of decelerations across all mice, n =22 trials, blue) and small decelerations (<20 percentile of decelerations from the distribution of decelerations across all mice, n = 108 trials, magenta). Shaded region, S.E.M. **C.** Same as (B) for acceleration. **D.** Same as (B) for spout licking. **E.** Left: ITI lick bout onset-aligned ∆F/F for the two fibers shown in Figure 5A-B across all onsets for the same session as B-D. Right: Mean lick bout triggered average ∆F/F for trials with large and small decelerations for trials in (E). Shaded region, S.E.M. **F-I.** Same as (B-E) for sessions post-learning. **J.** Box and whisker plots showing the mean ITI lick bout onset-evoked ΔF/F for each component (early peak, dip and late peak) for trials pre-learning with small and large decelerations across all fibers (blue dots). Box bounds indicate the 25th and 75th percentiles and the whiskers represent the minima and maxima. Significance calculated between large and small decelerations for each component across fibers using a two-tailed Wilcoxon rank-sum test. ***: $p < 0.001$. n.s.: not significant. **K.** Maps showing the difference in mean ITI lick bout onset-aligned ∆F/F with learning (post – pre) for the three signal components in mice expressing non-functional mutant ACh sensor (n = 3 mice). Each circle is a single fiber. Gray circles, no significant difference; Xs, no significant component. Significance of each component for each fiber was assessed using a two-tailed Wilcoxon rank-sum test, ($p < 0.01$).

References

1. Bolam, J.P., Wainer, B.H., and Smith, A.D. (1984). Characterization of cholinergic neurons in the rat neostriatum. A combination of choline acetyltransferase immunocytochemistry, Golgi-impregnation and electron microscopy. Neuroscience *12*, 711–718. https://doi.org/10.1016/0306-4522(84)90165-9.
2. Butcher, S.G., and Butcher, L.L. (1974). Origin and modulation of acetylcholine activity in the neostriatum. Brain Research *71*, 167–171. https://doi.org/10.1016/0006-8993(74)90202-9.
3. Aoki, S., Liu, A.W., Zucca, A., Zucca, S., and Wickens, J.R. (2015). Role of Striatal Cholinergic Interneurons in Set-Shifting in the Rat. J Neurosci *35*, 9424–9431. https://doi.org/10.1523/JNEUROSCI.0490-15.2015.
4. Aoki, S., Liu, A.W., Akamine, Y., Zucca, A., Zucca, S., and Wickens, J.R. (2018). Cholinergic interneurons in the rat striatum modulate substitution of habits. Eur J Neurosci *47*, 1194–1205. https://doi.org/10.1111/ejn.13820.
5. Bradfield, L.A., Bertran-Gonzalez, J., Chieng, B., and Balleine, B.W. (2013). The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. Neuron *79*, 153–166. https://doi.org/10.1016/j.neuron.2013.04.039.
6. Ragozzino, M.E. (2003). Acetylcholine actions in the dorsomedial striatum support the flexible shifting of response patterns. Neurobiol Learn Mem *80*, 257–267. https://doi.org/10.1016/s1074-7427(03)00077-7.
7. Ragozzino, M.E., Mohler, E.G., Prior, M., Palencia, C.A., and Rozman, S. (2009). Acetylcholine activity in selective striatal regions supports behavioral flexibility. Neurobiol Learn Mem *91*, 13–22. https://doi.org/10.1016/j.nlm.2008.09.008.
8. Ragozzino, M.E., Jih, J., and Tzavos, A. (2002). Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. Brain Res *953*, 205–214. https://doi.org/10.1016/s0006-8993(02)03287-0.
9. Fleming, W., Lee, J., Briones, B.A., Bolkan, S.S., and Witten, I.B. (2022). Cholinergic interneurons mediate cocaine extinction in male mice through plasticity across medium spiny neuron subtypes. Cell Reports *39*, 110874. https://doi.org/10.1016/j.celrep.2022.110874.
10. Collins, A.L., Aitken, T.J., Huang, I.-W., Shieh, C., Greenfield, V.Y., Monbouquette, H.G., Ostlund, S.B., and Wassum, K.M. (2019). Nucleus Accumbens Cholinergic Interneurons Oppose Cue-Motivated Behavior. Biol Psychiatry *86*, 388–396. https://doi.org/10.1016/j.biopsych.2019.02.014.
11. Gangal, H., Xie, X., Huang, Z., Cheng, Y., Wang, X., Lu, J., Zhuang, X., Essoh, A., Huang, Y., Chen, R., et al. (2023). Drug reinforcement impairs cognitive flexibility by inhibiting striatal cholinergic neurons. Nat Commun *14*, 3886. https://doi.org/10.1038/s41467-023-39623-x.
12. Kondabolu, K., Roberts, E.A., Bucklin, M., McCarthy, M.M., Kopell, N., and Han, X. (2016). Striatal cholinergic interneurons generate beta and gamma oscillations in the corticostriatal circuit and produce motor deficits. Proc Natl Acad Sci U S A *113*, E3159-3168. https://doi.org/10.1073/pnas.1605658113.
13. Huang, Z., Chen, R., Ho, M., Xie, X., Wang, X., and Wang, J. (2023). Dynamic Responses of Striatal Cholinergic Interneurons Control the Extinction and Updating of Goal-Directed Learning. Preprint, https://doi.org/10.1101/2023.11.09.566460 https://doi.org/10.1101/2023.11.09.566460.
14. Mallet, N., Leblois, A., Maurice, N., and Beurrier, C. (2019). Striatal Cholinergic Interneurons: How to Elucidate Their Function in Health and Disease. Front Pharmacol *10*, 1488. https://doi.org/10.3389/fphar.2019.01488.
15. Apicella, P. (2017). The role of the intrinsic cholinergic system of the striatum: What have we learned from TAN recordings in behaving animals? Neuroscience *360*, 81–94. https://doi.org/10.1016/j.neuroscience.2017.07.060.

16. Kimura, M., Rajkowski, J., and Evarts, E. (1984). Tonically discharging putamen neurons exhibit set-dependent responses. Proc. Natl. Acad. Sci. U.S.A. *81*, 4998–5001. https://doi.org/10.1073/pnas.81.15.4998.
17. Aosaki, T., Tsubokawa, H., Ishida, A., Watanabe, K., Graybiel, A.M., and Kimura, M. (1994). Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J. Neurosci. *14*, 3969–3984. https://doi.org/10.1523/JNEUROSCI.14-06-03969.1994.
18. Apicella, P., Legallet, E., and Trouche, E. (1997). Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states: Exp Brain Res *116*, 456–466. https://doi.org/10.1007/PL00005773.
19. Aosaki, T., Graybiel, A.M., and Kimura, M. (1994). Effect of the Nigrostriatal Dopamine System on Acquired Neural Responses in the Striatum of Behaving Monkeys. Science *265*, 412–415. https://doi.org/10.1126/science.8023166.
20. Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but Distinct Messages of Midbrain Dopamine and Striatal Tonically Active Neurons. Neuron *43*, 133–143. https://doi.org/10.1016/j.neuron.2004.06.012.
21. Gallo, E.F., Greenwald, J., Yeisley, J., Teboul, E., Martyniuk, K.M., Villarin, J.M., Li, Y., Javitch, J.A., Balsam, P.D., and Kellendonk, C. (2022). Dopamine D2 receptors modulate the cholinergic pause and inhibitory learning. Mol Psychiatry *27*, 1502–1514. https://doi.org/10.1038/s41380-021-01364-y.
22. Benhamou, L., Kehat, O., and Cohen, D. (2014). Firing Pattern Characteristics of Tonically Active Neurons in Rat Striatum: Context Dependent or Species Divergent? J. Neurosci. *34*, 2299–2304. https://doi.org/10.1523/JNEUROSCI.1798-13.2014.
23. Howe, M., Ridouh, I., Allegra Mascaro, A.L., Larios, A., Azcorra, M., and Dombeck, D.A. (2019). Coordination of rapid cholinergic and dopaminergic signaling in striatum during spontaneous movement. eLife *8*, e44903. https://doi.org/10.7554/eLife.44903.
24. Martyniuk, K.M., Torres-Herraez, A., Lowes, D.C., Rubinstein, M., Labouesse, M.A., and Kellendonk, C. (2022). Dopamine D2Rs coordinate cue-evoked changes in striatal acetylcholine levels. eLife *11*, e76111. https://doi.org/10.7554/eLife.76111.
25. Cragg, S.J. (2006). Meaningful silences: how dopamine listens to the ACh pause. Trends in Neurosciences *29*, 125–131. https://doi.org/10.1016/j.tins.2006.01.003.
26. Reynolds, J.N.J., Avvisati, R., Dodson, P.D., Fisher, S.D., Oswald, M.J., Wickens, J.R., and Zhang, Y.-F. (2022). Coincidence of cholinergic pauses, dopaminergic activation and depolarisation of spiny projection neurons drives synaptic plasticity in the striatum. Nat Commun *13*, 1296. https://doi.org/10.1038/s41467-022-28950-0.
27. Calabresi, P., Picconi, B., Tozzi, A., and Di Filippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. Trends Neurosci *30*, 211–219. https://doi.org/10.1016/j.tins.2007.03.001.
28. Yarom, O., and Cohen, D. (2011). Putative Cholinergic Interneurons in the Ventral and Dorsal Regions of the Striatum Have Distinct Roles in a Two Choice Alternative Association Task. Front. Syst. Neurosci. *5*. https://doi.org/10.3389/fnsys.2011.00036.
29. Atallah, H.E., McCool, A.D., Howe, M.W., and Graybiel, A.M. (2014). Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. Neuron *82*, 1145–1156. https://doi.org/10.1016/j.neuron.2014.04.021.
30. Mohebi, A., Collins, V.L., and Berke, J.D. (2023). Accumbens cholinergic interneurons dynamically promote dopamine release and enable motivation. eLife *12*, e85011. https://doi.org/10.7554/eLife.85011.
31. Marche, K., Martel, A.-C., and Apicella, P. (2017). Differences between Dorsal and Ventral Striatum in the Sensitivity of Tonically Active Neurons to Rewarding Events. Front Syst Neurosci *11*, 52. https://doi.org/10.3389/fnsys.2017.00052.
32. Jing, M., Li, Y., Zeng, J., Huang, P., Skirzewski, M., Kljakic, O., Peng, W., Qian, T., Tan, K.,

Zou, J., et al. (2020). An optimized acetylcholine sensor for monitoring in vivo cholinergic activity. Nat Methods *17*, 1139–1146. https://doi.org/10.1038/s41592-020-0953-2.
33. Engelbrecht, C.J., Göbel, W., and Helmchen, F. (2009). Enhanced fluorescence signal in nonlinear microscopy through supplementary fiber-optic light collection. Opt. Express, OE *17*, 6421–6435. https://doi.org/10.1364/OE.17.006421.
34. Pisanello, M., Pisano, F., Hyun, M., Maglie, E., Balena, A., De Vittorio, M., Sabatini, B.L., and Pisanello, F. (2019). The Three-Dimensional Signal Collection Field for Fiber Photometry in Brain Tissue. Front. Neurosci. *13*. https://doi.org/10.3389/fnins.2019.00082.
35. Perkins, L.N., Semu, D., Shen, J., Boas, D.A., and Gardner, T.J. (2018). High-density microfibers as a potential optical interface to reach deep brain regions. J. Neural Eng. *15*, 066002. https://doi.org/10.1088/1741-2552/aadbb2.
36. Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., et al. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell *181*, 936-953.e20. https://doi.org/10.1016/j.cell.2020.04.007.
37. Vu, M.-A.T., Brown, E.H., Wen, M.J., Noggle, C.A., Zhang, Z., Monk, K.J., Bouabid, S., Mroz, L., Graham, B.M., Zhuo, Y., et al. (2024). Targeted micro-fiber arrays for measuring and manipulating localized multi-scale neural dynamics over large, deep brain volumes during behavior. Neuron *112*, 909-923.e9. https://doi.org/10.1016/j.neuron.2023.12.011.
38. Krok, A.C., Maltese, M., Mistry, P., Miao, X., Li, Y., and Tritsch, N.X. (2023). Intrinsic dopamine and acetylcholine dynamics in the striatum of mice. Nature *621*, 543–549. https://doi.org/10.1038/s41586-023-05995-9.
39. Zhang, W.-T., Chao, T.-H.H., Yang, Y., Wang, T.-W., Lee, S.-H., Oyarzabal, E.A., Zhou, J., Nonneman, R., Pegard, N.C., Zhu, H., et al. (2022). Spectral fiber photometry derives hemoglobin concentration changes for accurate measurement of fluorescent sensor activity. Cell Reports Methods *2*, 100243. https://doi.org/10.1016/j.crmeth.2022.100243.
40. Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L., and Tank, D.W. (2010). Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. Nat Neurosci *13*, 1433–1440. https://doi.org/10.1038/nn.2648.
41. Apicella, P., Deffains, M., Ravel, S., and Legallet, E. (2009). Tonically active neurons in the striatum differentiate between delivery and omission of expected reward in a probabilistic task context. European Journal of Neuroscience *30*, 515–526. https://doi.org/10.1111/j.1460-9568.2009.06872.x.
42. Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and Bergman, H. (2008). Midbrain Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning Trials. J. Neurosci. *28*, 11673–11684. https://doi.org/10.1523/JNEUROSCI.3839-08.2008.
43. Ding, J.B., Guzman, J.N., Peterson, J.D., Goldberg, J.A., and Surmeier, D.J. (2010). Thalamic gating of corticostriatal signaling by cholinergic interneurons. Neuron *67*, 294–307. https://doi.org/10.1016/j.neuron.2010.06.017.
44. Kharkwal, G., Radl, D., Lewis, R., and Borrelli, E. (2016). Dopamine D2 receptors in striatal output neurons enable the psychomotor effects of cocaine. Proceedings of the National Academy of Sciences *113*, 11609–11614. https://doi.org/10.1073/pnas.1608362113.
45. Zhang, Y.-F., Reynolds, J.N.J., and Cragg, S.J. (2018). Pauses in Cholinergic Interneuron Activity Are Driven by Excitatory Input and Delayed Rectification, with Dopamine Modulation. Neuron *98*, 918-925.e3. https://doi.org/10.1016/j.neuron.2018.04.027.
46. Hunnicutt, B.J., Jongbloets, B.C., Birdsong, W.T., Gertz, K.J., Zhong, H., and Mao, T. (2016). A comprehensive excitatory input map of the striatum reveals novel functional organization. eLife *5*, e19103. https://doi.org/10.7554/eLife.19103.
47. Hintiryan, H., Foster, N.N., Bowman, I., Bay, M., Song, M.Y., Gou, L., Yamashita, S., Bienkowski, M.S., Zingg, B., Zhu, M., et al. (2016). The mouse cortico-striatal projectome. Nat Neurosci *19*, 1100–1114. https://doi.org/10.1038/nn.4332.

48. Schultz, W., Dayan, P., and Montague, P.R. (1997). A Neural Substrate of Prediction and Reward. Science *275*, 1593–1599. https://doi.org/10.1126/science.275.5306.1593.
49. van Elzelingen, W., Goedhoop, J., Warnaar, P., Denys, D., Arbab, T., and Willuhn, I. (2022). A unidirectional but not uniform striatal landscape of dopamine signaling for motivational stimuli. Proceedings of the National Academy of Sciences *119*, e2117270119. https://doi.org/10.1073/pnas.2117270119.
50. Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci *10*, 1020–1028. https://doi.org/10.1038/nn1923.
51. Gritton, H.J., Howe, W.M., Romano, M.F., DiFeliceantonio, A.G., Kramer, M.A., Saligrama, V., Bucklin, M.E., Zemel, D., and Han, X. (2019). Unique contributions of parvalbumin and cholinergic interneurons in organizing striatal networks during movement. Nat Neurosci *22*, 586–597. https://doi.org/10.1038/s41593-019-0341-3.
52. Shroff, S.N., Lowet, E., Sridhar, S., Gritton, H.J., Abumuaileq, M., Tseng, H.-A., Cheung, C., Zhou, S.L., Kondabolu, K., and Han, X. (2023). Striatal cholinergic interneuron membrane voltage tracks locomotor rhythms in mice. Nat Commun *14*, 3802. https://doi.org/10.1038/s41467-023-39497-z.
53. Chantranupong, L., Beron, C.C., Zimmer, J.A., Wen, M.J., Wang, W., and Sabatini, B.L. (2023). Dopamine and glutamate regulate striatal acetylcholine in decision-making. Nature *621*, 577–585. https://doi.org/10.1038/s41586-023-06492-9.
54. Kim, T., Capps, R.A., Hamade, K.C., Barnett, W.H., Todorov, D.I., Latash, E.M., Markin, S.N., Rybak, I.A., and Molkov, Y.I. (2019). The Functional Role of Striatal Cholinergic Interneurons in Reinforcement Learning From Computational Perspective. Front. Neural Circuits *13*. https://doi.org/10.3389/fncir.2019.00010.
55. Deffains, M., and Bergman, H. (2015). Striatal cholinergic interneurons and cortico-striatal synaptic plasticity in health and disease. Movement Disorders *30*, 1014–1025. https://doi.org/10.1002/mds.26300.
56. Zhang, Y.-F., and Cragg, S.J. (2017). Pauses in Striatal Cholinergic Interneurons: What is Revealed by Their Common Themes and Variations? Front. Syst. Neurosci. *11*. https://doi.org/10.3389/fnsys.2017.00080.
57. Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci *1*, 304–309. https://doi.org/10.1038/1124.
58. Steinberg, E.E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K., and Janak, P.H. (2013). A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci *16*, 966–973. https://doi.org/10.1038/nn.3413.
59. Chuhma, N., Mingote, S., Moore, H., and Rayport, S. (2014). Dopamine Neurons Control Striatal Cholinergic Neurons via Regionally Heterogeneous Dopamine and Glutamate Signaling. Neuron *81*, 901–912. https://doi.org/10.1016/j.neuron.2013.12.027.
60. Straub, C., Tritsch, N.X., Hagan, N.A., Gu, C., and Sabatini, B.L. (2014). Multiphasic Modulation of Cholinergic Interneurons by Nigrostriatal Afferents. Journal of Neuroscience *34*, 8557–8569. https://doi.org/10.1523/JNEUROSCI.0589-14.2014.
61. Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.-H., Folk, R.W., Broussard, G.J., Liang, R., Jang, M.J., et al. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. Science *360*, eaat4422. https://doi.org/10.1126/science.aat4422.
62. Marvin, J.S., Scholl, B., Wilson, D.E., Podgorski, K., Kazemipour, A., Müller, J.A., Schoch, S., Quiroz, F.J.U., Rebola, N., Bao, H., et al. (2018). Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. Nat Methods *15*, 936–939. https://doi.org/10.1038/s41592-018-0171-3.
63. Reynolds, J.N., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. Nature *413*, 67–70. https://doi.org/10.1038/35092560.

64. Wang, Z., Kai, L., Day, M., Ronesi, J., Yin, H.H., Ding, J., Tkatch, T., Lovinger, D.M., and Surmeier, D.J. (2006). Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. Neuron *50*, 443–452. https://doi.org/10.1016/j.neuron.2006.04.010.
65. Goldberg, J.A., Ding, J.B., and Surmeier, D.J. (2012). Muscarinic modulation of striatal function and circuitry. Handb Exp Pharmacol, 223–241. https://doi.org/10.1007/978-3-642-23274-9_10.
66. Duhne, M., Mohebi, A., Kim, K., Pelattini, L., and Berke, J. (2024). A mismatch between striatal cholinergic pauses and dopaminergic reward prediction errors. Preprint, https://doi.org/10.1101/2024.05.09.593336 https://doi.org/10.1101/2024.05.09.593336.
67. Girasole, A.E., and Nelson, A.B. (2015). Probing striatal microcircuitry to understand the functional role of cholinergic interneurons. Mov Disord *30*, 1306–1318. https://doi.org/10.1002/mds.26340.
68. Shen, W., Plotkin, J.L., Francardo, V., Ko, W.K.D., Xie, Z., Li, Q., Fieblinger, T., Wess, J., Neubig, R.R., Lindsley, C.W., et al. (2015). M4 Muscarinic Receptor Signaling Ameliorates Striatal Plasticity Deficits in Models of L-DOPA-Induced Dyskinesia. Neuron *88*, 762–773. https://doi.org/10.1016/j.neuron.2015.10.039.
69. Nair, A.G., Castro, L.R.V., El Khoury, M., Gorgievski, V., Giros, B., Tzavara, E.T., Hellgren-Kotaleski, J., and Vincent, P. (2019). The high efficacy of muscarinic M4 receptor in D1 medium spiny neurons reverses striatal hyperdopaminergia. Neuropharmacology *146*, 74–83. https://doi.org/10.1016/j.neuropharm.2018.11.029.
70. Lee, S.J., Lodder, B., Chen, Y., Patriarchi, T., Tian, L., and Sabatini, B.L. (2021). Cell-type-specific asynchronous modulation of PKA by dopamine in learning. Nature *590*, 451–456. https://doi.org/10.1038/s41586-020-03050-5.
71. Sergejeva, M., Papp, E.A., Bakker, R., Gaudnek, M.A., Okamura-Oho, Y., Boline, J., Bjaalie, J.G., and Hess, A. (2015). Anatomical landmarks for registration of experimental image data to volumetric rodent brain atlasing templates. J Neurosci Methods *240*, 161–169. https://doi.org/10.1016/j.jneumeth.2014.11.005.
72. Zhang, Y., Zhao, S., Rodriguez, E., Takatoh, J., Han, B.-X., Zhou, X., and Wang, F. (2015). Identifying local and descending inputs for primary sensory neurons. J Clin Invest *125*, 3782–3794. https://doi.org/10.1172/JCI81156.
73. Miri, A., Daie, K., Arrenberg, A.B., Baier, H., Aksay, E., and Tank, D.W. (2011). Spatial gradients and multidimensional dynamics in a neural integrator circuit. Nat Neurosci *14*, 1150–1159. https://doi.org/10.1038/nn.2888.