TITLE:

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

AUTHORS

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,*]

AFFILIATIONS

[1]Department of Psychological & Brain Sciences, Boston University, Boston, MA, USA

[*]Lead contact; Correspondence: mwhowe@bu.edu

ABSTRACT

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

KEYWORDS

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

INTRODUCTION

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

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

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

heterogeneities in ACh signaling 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. Standard electrode implants measure spiking from only a few unique CINs per subject, typically target limited striatal regions (< 4), and without optical tagging can suffer from ambiguities in accurate CIN identification. Fiber photometry has been used to optically measure ACh release, but similarly has been limited to 1-2 striatal regions per experiment. Therefore, it is poorly understood how and whether ACh release dynamics differ across the striatum volume to support distinct functions in associative learning and extinction.

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

RESULTS

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

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[44] (50μm diameter) distributed uniformly across the 3-dimensional striatum volume (Fig. 1a-e). Each fiber collected fluorescence from post-synaptic neurons expressing the genetically encoded ACh sensor, GRAB-ACh3.0[45] (Fig. 1a). Based on previous empirical measurements and light scattering models[46,47], we estimate that our fibers collect fluorescence from a tapered collection volume extending approximately 100μm axially and 25μm radially from each fiber tip (see Vu et al, 2024 for details[48]). The minimum separation of fiber tips in our arrays was 220μm radially and 250μm axially ensuring no (or very minimal) overlap in the collection volumes for each fiber. 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)[48,49]. Rapid, bi-directional changes in fluorescence, a proxy for changes in ACh release and synchronous firing of cholinergic interneurons[50], could be reliably detected across striatum locations (Fig. 1c). 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 (Figs. 1f, g). However, all three

signal components were not present in every region (Fig. 1g). 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 statistically significant changes in unpredicted reward mean $\Delta F/F$ across days of learning (1.6%, 5.7% and 5.7% of the 295 total 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. 1f). Thus, our optical approach enabled us to investigate the evolution of simultaneous, spatiotemporal patterns of ACh release across the striatum during learning.

To control for potential signal contamination by hemodynamic or motion artifacts, we conducted measurements with quasi-simultaneous 405nm illumination, the approximate isosbestic point of the sensor[45] and in mice expressing a non-functional null version of the ACh3.0 sensor (Fig. 1h, i). 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, Fig. 1i). In mice expressing the null sensor, we found that some small artifacts were present with 470nm illumination which were not detected with 405nm illumination, and overall, the artifactual fluorescence changes measured with 405nm illumination were only weakly correlated with 470nm illumination (Fig. 1h). 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[51]. 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.

Spatially organized, opposing 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[52] 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 (p = 0.0078, two-tailed Wilcoxon signed rank test, Fig. 2b-d), and they decreased non-contingent, 'spontaneous' licking during the ITI period (p = 0.0273, two-tailed Wilcoxon signed rank test, Fig. 2e). A lick index was computed, reflecting the relative proportion of anticipatory licking within the cue period relative to the ITI, to assess learning of the cue-reward association on each session (Fig. 2c).

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

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

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

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

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

Emergence of spatially organized ACh signals during extinction

We next tested how and where cue-evoked ACh release changes as cues lose their predictive value 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 lower probability cue, indicating successful cue-specific extinction learning, at which point, reward probabilities for each cue were reversed (Fig. 4b-d). ACh early peaks predominantly decreased relative to the post learning phase for the low probability cue (light: 36.27%, tone: 40.34% of total fibers, Fig.4e,f; Fig S4c) and dips became smaller (light: 22.03%, tone: 29.5% of total fibers, Fig. 4g,h; Fig S4c). The direction of these changes represented a reversal of the changes during initial learning (Fig. 2, Fig.S2). However, unlike the early peaks and dips, ACh late peaks primarily increased for the low probability cue relative to late learning for both cues (light: 27.46%, tone: 20.68% of total fibers, Fig. 4g,h; Fig S4c). Differences were consistent for the initial extinction and reversal phases, so data was combined for all analyses. These results indicate that the degradation of cue-reward associations is accompanied by a reversal of learned dynamics for early peak and dip components but an emergence of new ACh increases for the late peak component.

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

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

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

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

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

ACh signaling has been previously linked to changes in locomotion [39,50,58,59], so we asked whether changes we observed across learning phases 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. During initial Pavlovian learning, mice decelerated significantly more, on average, after cue onset and less after reward delivery as they learned the cue-reward associations (Fig. S7a,b, $p > 0.05$ for peak deceleration and $p<0.05$ for velocity changes after cue onset, Freidman test; $p <0.05$ for peak deceleration and $p>0.05$ for velocity after reward delivery, Wilcoxon rank-sum test). Differences were not present over learning for spontaneous ITI lick bouts (Fig. S7c, $p>0.05$, Wilcoxon rank-sum test) or during extinction for high and low probability cues or relative to the late learning phase (Fig. S7a, $p > 0.05$, Freidman test). To address the possibility that locomotion deceleration signals may contribute to the region-specific ACh changes that we report with initial learning (Figs. 2 and 3; Figs. S1-3), we first examined ACh release to spontaneous decelerations in the ITI period with similar magnitude to decelerations at cues and rewards (Fig. S7d-g). At some locations, primarily in the dorsal striatum, we observed bi-phasic peaks and dips in ACh release associated with ITI decelerations (Fig. S7d). However, these peaks and dips in $\Delta F/F$ were present across fewer fibers and were much lower in magnitude relative to changes to cues and rewards with learning (Fig. S7e-g). Most fibers with large cue and reward changes with learning did not have any significant ITI deceleration response (36.15% and 23.76% of fibers with significant peak and dip changes at cue had significant maxima and minima for ITI decelerations respectively; 50% and 25.43% for fibers with peaks and dips at reward respectively; Fig. S7f,g). In addition, change magnitudes at cue and reward with learning were not correlated with magnitudes of peaks and dips to spontaneous decelerations across fibers (Fig. S7l,m; Table 1). For most fibers, 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 decelerations at cue and reward (Figs. 2, 3; Fig.S7h,i). Moreover, ACh release to cues, rewards, and ITI licks were mostly insensitive to trial-by-trial variations in the size of the associated deceleration (Fig. S7h-k). ACh peaks and dips for only a small fraction of fibers with significant response components were significantly correlated with trial by trial deceleration for cues (7.4% of 270, 7.1% of /224, and 1.1% of /180 fibers for early peak, dip, and late peak of light cue respectively; 8.3% of 262, 5.1% of 176, and 3.9% of 76 locations for early peak, dip, and late peak of tone cue respectively; Pearson's correlation, $p<0.01$) and rewards (19.1% of 287, 23.5% of 276, and 6.1% of 229 locations for

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

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

Our ACh release measurements identified transient increases in cue-evoked ACh release in the aDS (late peaks), consistent with negative RPEs, which may contribute to downshifting the predictive value of cues and the expression of appetitive behaviors, such as licking, when reward probabilities change (Figs. 4 and 5). 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. Targeting was informed by the distribution of late peak changes for the light cue (Fig. 4j), so injections were shifted slightly to favor the dorsal medial striatum (aDMS, see Methods, Fig. 6a-c). TelC expression eliminates synaptic vesicle exocytosis in cholinergic interneurons, and consequently blocks ACh release[60] (Fig. 6c). ACh suppression did not affect the mice ability to learn the cue-reward association, as both TelC and control mice increased the frequency of licking in the cue period over learning (Two-way repeated measure ANOVA , p = 0.02 for Telc and p = 0.0001 for control, Fig. 6d). 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[19,61–63] (Fig. 2i,k; Fig. S2c,d). 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. 6d,f). 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. 6d,f). TelC mice also had slightly higher licking rates during the ITI and cue periods than the control mice at all phases of learning, perhaps indicating a general difference in suppression of spontaneous licking (Fig. 6d,e). Although the manipulation was not temporally specific, these data indicate that intact ACh release in the aDMS contributes to the extinction of learned Pavlovian responses following a downshift in cue associated reward probabilities.

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

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[34,35,55,64]. If present, these DA dips may disinhibit CINs through reduced tonic activation of inhibitory D2 receptors[13,60,65,66]. We tested whether DA signals in the aDS are consistent with a potential role in facilitating the observed elevations in ACh signaling. We optically measured DA release dynamics with dLight1.3b[67] in the same regions of the aDS where we observed the increases in ACh peaks to unrewarded licks and down-shifted cues (Figs. 4,5,7a). Consistent with prior work, DA release rapidly increased to conditioned cues following learning (Fig. 7b,c). Following partial

extinction (Fig. 4a), dips below baseline emerged after short latency peaks, selectively for the low probability cue (Fig. 7b-d). The average latency to the trough of DA dips (median to trough = 0.63s) preceded the average latency of the ACh late peaks (median = 0.8s) recorded in the same aDS region in the ACh3.0 expressing group (Fig. 7e, p = 0.02, Two-tailed rank-sum test). Like the ACh late peaks, dips in cue-evoked aDS DA 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). Like the ACh3.0 expressing mice, there were no differences in the dLight1.3b group in cue-related decelerations between the post learning and extinction phases, so DA dips could not be explained by locomotion changes (Fig. 7g). We then asked whether DA release for spontaneous ITI licks also exhibited learning related dips reflecting shifts in negative RPE signaling with learning. DA release in the aDS increased for spontaneous licks, rising just prior to spout contact (Fig. 7h,i). 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). Like the cue responses, DA dips had a shorter average latency (median = 0.33s) relative to the ACh late peaks (median = 0.44s, p= 3.33x10^{-6}, two-tailed Wilcoxon rank-sum test, Fig. 7k). These results indicate that aDS DA release reflects putative negative prediction errors in dips to extinction cues and unrewarded ITI licks, which are opposite in polarity and precede the average latency of ACh late peaks in the same region.

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

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[68] selectively in aDS cholinergic interneurons of ChAT-cre mice to measure rapid changes in glutamate release (GluCIN) during partial extinction (Fig. 8a-c). 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. 8d,e). Consistent with the relative absence of strong ACh release peaks in the aDS to the conditioned tone (Figs. S1d, S2d,f), 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. Following light cue partial extinction (from 100 to 20% reward probability), there was no significant change in this bi-phasic profile (Fig. 8e, early peak: p = 0.96, two-tailed Wilcoxon rank-sum test; late peak: p = 0.34, two-tailed Wilcoxon rank-sum test) indicating that an increase in glutamate release onto the CINs was not responsible for the increase of aDS ACh release during partial extinction (Fig. 4). Despite this, the latency of the slower component of the GluCIN release aligned, on average, with the latency of the ACh late peak increase following extinction (median latency to Glu late 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 reveal a topographic organization in the evolution of ACh release across timescales ranging from 100s of milliseconds to days as Pavlovian associations are learned then broken. Most notably, we identified a specific region of the anterior dorsal striatum (aDS) with changes in ACh release over learning and extinction consistent with positive and negative RPE encoding for cues, rewards, and unrewarded consummatory actions. Changes were inverse in polarity to those widely reported for DA release and RPE encoding was preferentially expressed in long latency components (dips and late peaks) of the multi-phasic ACh release profiles. Some changes in early peaks consistent with positive RPE encoding similar to DA were observed[22,39], but were more sparse, concentrated ventrally, and were primarily restricted to the cue period. We hypothesize that positive prediction error encoding in ACh dips defines plasticity windows over time and space which facilitate initial formation of Pavlovian associations. In contrast, negative prediction error encoding in ACh late peak elevations may play an active role in erasing learned associations which are no longer valid, resulting in down-regulated behavioral responding to previously valued cues. These signals may support a general role for striatal ACh in forms of behavioral flexibility identified in previous manipulation studies. While our study identified the aDS as a potential locus for downshifting learned cue-reward associations, different forms of flexibility may recruit changes in ACh in distinct striatal regions. Indeed, one study observed a correlation between ACh transients in the medial nucleus accumbens shell and cocaine place preference extinction[29] and another observed ACh increases during extinction of lever pressing behavior in the dorsal medial striatum[33]. Further, ACh manipulations can result in diverse, sometimes conflicting, effects across task paradigms and regions. Future surveys of large-scale ACh release will be necessary to clarify the specific ACh dynamics across striatal regions associated with different forms of behavioral flexibility.

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

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

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

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

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

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

Extinction has been proposed to occur both through an 'unlearning' process involving reversal of plasticity occurring during associative acquisition and through new inhibitory learning[64,76]. The ACh peaks in the aDS are well positioned to promote both of these processes in parallel through their differential effects of synaptic plasticity on striatal projection neuron cell types. Pavlovian associations are believed to involve the strengthening of synapses onto D1 receptor expressing direct pathway neurons (dSPNs) driven by phasic elevations in DA to unpredicted rewards and reward-associated cues (e.g. positive RPEs)[21,55,74]. Studies *in-vitro* have found that M4 muscarinic ACh receptors, expressed preferentially on direct pathway spiny projection neurons (dSPNs), promote long term synaptic depression[77,78]. Thus, the increases in cue evoked ACh during extinction could weaken the previously potentiated synapses onto dSPNs. Importantly, while DA dips encoding negative RPEs have been proposed to contribute to extinction, it is unclear how DA dips alone could weaken 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[79] . In contrast, CINs express the high affinity inhibitory D2 DA receptor, so DA dips promote rapid ACh increases through disinhibition[60,79]. Therefore, disinhibition of CINs may be a mechanism to convert negative prediction errors encoded by DA dips into ACh signals that can drive de-potentiation of dSPN synapses and 'unlearning' of invalid cue-reward associations. In parallel, the aDS ACh elevations (and DA dips) during extinction may elevate excitability (relative to dSPNs) and strengthen synapses of D2 expressing indirect pathway neurons (iSPNs) via M1 receptors[3,4,80]. This may result in new inhibitory learning, where iSPNs become more strongly activated to the extinction cue to actively suppress appetitive responding.

METHODS

Mice

For all experiments adult male and female (n = 33 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=18, 10 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

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

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

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

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

Stereotaxic viral injections and chronic optical fiber 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 (WZ Biosciences[45]), 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[45] (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 DA release using multi-fiber photometry, AAV5-CAG-dlight1.3b[67] (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 release 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 (Addgene, #135391,[84], 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 anterior dorsal medial striatum 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. To validate that TelC expression in the aDS leads to a reduction in ACh release, two ChAT-cre mice were bilaterally injected in the aDS using the same coordinates with ssAAV-5/2-hSyn1-dlox-TeTxLC_2A_NLS_dTomato(rev)-dlox-WRPE-hGHp diluted 1:1 in PBS. Additionally, the left hemisphere of each mouse was co-injected with

AAV9-hSyn-ACh3.0 diluted 1:2 in PBS. 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[68] (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 medial region of the aDS (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[52] (Smoothfoam, 8in diameter, Fig. 1b). The behavioral setup has been described in detail previously[48]. 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, 2a and 4a). 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). 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[48]. 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[48] . 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. 1a). 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[52,85] 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[48]. 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, Fig. 1h,i). 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 (Fig.1f).

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:

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

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 (2.9±0.62), while sessions where mean lick indices of both cues were significantly higher than the first session were considered post-learning sessions (3.27±0.5) (Fig. 2c,d). Sessions where only one cue had a significantly elevated lick index were omitted from analysis (2.90±0.75 sessions). 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. 2d) 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 (Figs. 5 and 7) 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[48], 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, $\widehat{\beta}$ are the estimated correlation coefficients generated by the GLM. Non-significant $\widehat{\beta}$, 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. S7). To identify decelerations occurring to cues, rewards, or lick bout onsets, deceleration periods were first defined in peri-event windows as consecutive bins with negative acceleration values. Peak deceleration (Fig. S7a-c; Fig. 7g) 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 (Fig. S7a-c; Fig. 7g) were defined as the maximum change of velocity following the event onset, relative to pre-event. Total velocity change and peak decelerations were calculated from cue onset to 0.6s after, and from -0.2 to 0.6s relative to reward consumption (first spout lick after delivery). 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 (Fig. S7h-k). For analyzing the contribution of signals related to spontaneous, non-task related decelerations (Fig. S7d-g), decelerations were first identified within ITI periods, excluding all times around spontaneous licks and unpredicted reward deliveries. Periods of continuous negative acceleration were identified and the troughs were included in triggered average analysis if they were lower than -1.5m/s^2 , and there was a total velocity change over the deceleration period of at least 20 cm/s. These thresholds were chosen conservatively to ensure that ITI decelerations matched (or exceeded) the average deceleration magnitudes observed in the task (Fig S7a-c). ACh ΔF/F signals were aligned to the deceleration trough, and maximum and minimum values of the mean trough triggered average were calculated from a ± 1s window around the peak. Statistical significance of minima and maxima was determined by comparison to the 95% confidence intervals of a bootstrap-generated random sampling distribution (2000 iterations). To evaluate the potential influence of decelerations occurring at the cue or reward on learning-related changes, correlations (Pearson's) were calculated between significant ITI deceleration minima or maxima and the magnitude of the mean ΔF/F for different

signal components during the cue and reward periods, limited to fibers with both significant deceleration and task-related changes for each component (Fig. S7l,m; Table 1).

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

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 peri-event windows (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. S1a,b; Fig. S6a). Based on the multi-phasic timing distributions from the histograms, windows were defined for each event to define significant early peak, dip, and late peak components (Fig. S1a,b). 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 an unpaired Wilcoxon rank sum test. *P*-values <0.01 were considered statistically significant, except for the difference in mean ITI lick bout-evoked ΔF/F, where p-values <0.025 were considered significant. Signal changes consistent with RPE signaling were determined for each signal component and fiber across pairs of events (Figs. 3d,5g; Figs. S3e,S6f). Putative positive RPE encoding was defined as elevations in ΔF/F (larger peaks or deeper dips) with learning to the cue and decreases to the reward. Negative RPE encoding was defined as elevations to the cue with extinction and decreases to unrewarded ITI licks with learning.

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

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

Spatial gradient analysis

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

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

where y_{ijk} is the amplitude of the ΔF/F differences, β_0 is the intercept, β_{AP}, β_{ML}, β_{DV} are the fixed effect coefficients for the AP, ML and DV coordinates respectively, u_{mouse}, $u_{AP,\,mouse}$, $u_{ML,\,mouse}$, $u_{DV,\,mouse}$ are random effect intercept and coefficients for AP, ML and DV coordinates, for individual mice, and ϵ_{ijk} is the residual error term. [β_{AP}, β_{ML}, β_{DV}] defined the striatal axes of maximal variation. The sign of the spatial coefficients was set so that the direction of the predominant signal change for each component was positive (e.g. larger dips or larger peaks were both positive), so that the sign of the coefficient indicated the direction of the spatial gradient (rather than the sign of the change) for comparison across components that changed in opposite directions.

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

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

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

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

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. 4m-p and 7f), 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.

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 DV>-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. 7k). 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. 7f).

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 Aligning Science Across Parkinson's (ASAP, ASAP-020370) through the Michael J. Fox Foundation for Parkinson's Research (MJFF), Stanley Fahn Young Investigator Award (PF-SF-JFA-836662), Klingenstein-Simons's Foundation fellowship, Whitehall Foundation Fellowship, National Institute of Mental Health (R01 MH125835) 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 and Eleanor Brown 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, MWH; 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.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Data availability

Source data are available upon request from the corresponding author and will be made open access before publishing.

Code availability

Code is available upon request from the corresponding author and will be made open access before publishing.

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. 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.
4. Lim, S.A.O., Kang, U.J., and McGehee, D.S. (2014). Striatal cholinergic interneuron regulation and circuit effects. Front. Synaptic Neurosci. *6*. https://doi.org/10.3389/fnsyn.2014.00022.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.

16. 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.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.
22. Duhne, M., Mohebi, A., Kim, K., Pelattini, L., and Berke, J.D. (2024). A mismatch between striatal cholinergic pauses and dopaminergic reward prediction errors. Proc. Natl. Acad. Sci. U.S.A. *121*, e2410828121. https://doi.org/10.1073/pnas.2410828121.
23. 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.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. 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.
30. 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.
31. 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.

32. 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.
33. Huang, Z., Chen, R., Ho, M., Xie, X., Gangal, H., Wang, X., and Wang, J. (2024). Dynamic responses of striatal cholinergic interneurons control behavioral flexibility. Sci. Adv. *10*, eadn2446. https://doi.org/10.1126/sciadv.adn2446.
34. 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.
35. 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.
36. Lak, A., Stauffer, W.R., and Schultz, W. (2016). Dopamine neurons learn relative chosen value from probabilistic rewards. eLife *5*, e18044. https://doi.org/10.7554/eLife.18044.
37. Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C.K., Hassabis, D., Munos, R., and Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. Nature *577*, 671–675. https://doi.org/10.1038/s41586-019-1924-6.
38. 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.
39. 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.
40. 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.
41. 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.
42. 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.
43. 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.
44. 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.
45. 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.
46. 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*, 82. https://doi.org/10.3389/fnins.2019.00082.
47. 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.
48. 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.

49. 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.

50. 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.

51. 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.

52. 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.

53. 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.

54. 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.

55. 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.

56. 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.

57. 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.

58. 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.

59. 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.

60. 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.

61. 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.

62. 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.

63. 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.

64. Pan, W.-X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2008). Tripartite mechanism of extinction suggested by dopamine neuron activity and temporal difference model. J Neurosci *28*, 9619–9631. https://doi.org/10.1523/JNEUROSCI.0255-08.2008.
65. 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.
66. 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.
67. 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.
68. 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.
69. 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.
70. Azcorra, M., Gaertner, Z., Davidson, C., He, Q., Kim, H., Nagappan, S., Hayes, C.K., Ramakrishnan, C., Fenno, L., Kim, Y.S., et al. (2023). Unique functional responses differentially map onto genetic subtypes of dopamine neurons. Nat Neurosci *26*, 1762–1774. https://doi.org/10.1038/s41593-023-01401-9.
71. Lerner, T.N., Shilyansky, C., Davidson, T.J., Evans, K.E., Beier, K.T., Zalocusky, K.A., Crow, A.K., Malenka, R.C., Luo, L., Tomer, R., et al. (2015). Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. Cell *162*, 635–647. https://doi.org/10.1016/j.cell.2015.07.014.
72. Howe, M.W., and Dombeck, D.A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature *535*, 505–510. https://doi.org/10.1038/nature18942.
73. Tsutsui-Kimura, I., Matsumoto, H., Akiti, K., Yamada, M.M., Uchida, N., and Watabe-Uchida, M. (2020). Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. eLife *9*, e62390. https://doi.org/10.7554/eLife.62390.
74. 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.
75. 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.
76. Delamater, A.R., and Westbrook, R.F. (2014). Psychological and neural mechanisms of experimental extinction: A selective review. Neurobiology of Learning and Memory *108*, 38–51. https://doi.org/10.1016/j.nlm.2013.09.016.
77. 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.
78. 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.

79. 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.

80. Shen, W., Hamilton, S.E., Nathanson, N.M., and Surmeier, D.J. (2005). Cholinergic Suppression of KCNQ Channel Currents Enhances Excitability of Striatal Medium Spiny Neurons. J. Neurosci. *25*, 7449–7458. https://doi.org/10.1523/JNEUROSCI.1381-05.2005.

81. 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.

82. 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.

83. 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.

84. 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.

85. 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.

Fig. 1

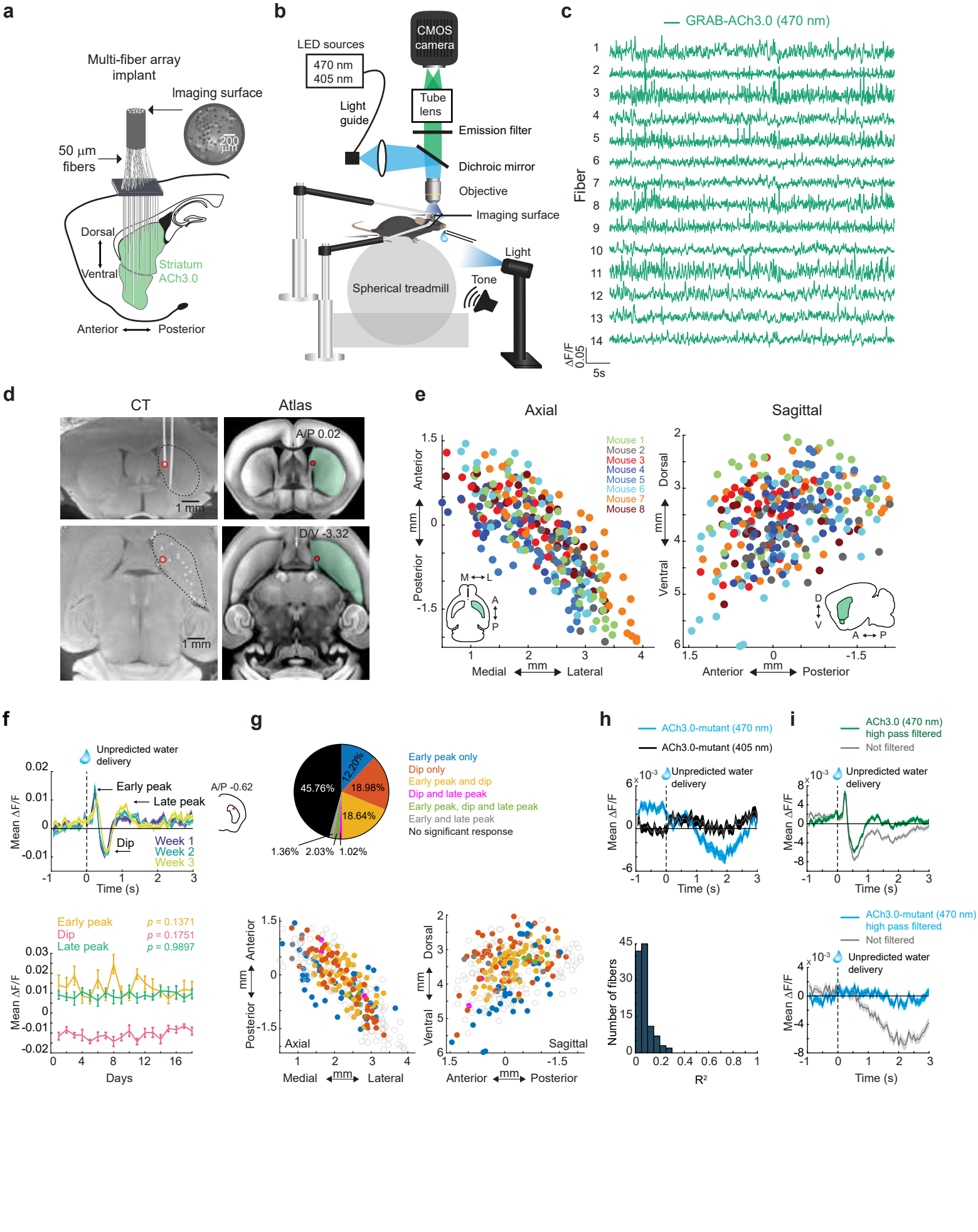


FIGURE TITLES AND LEGENDS

Fig. 1: 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 (ΔF/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** Top: mean ΔF/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 (Fig. 2). Inset indicates fiber location in the coronal plane. Shaded regions, S.E.M. Bottom: The mean peak (early/late) or dip ΔF/F for unpredicted reward deliveries across 18 consecutive sessions of training for the example at top. 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. **g** Top: percentage of the total fibers (295 fibers) across all mice with significant ($p < 0.01$, two-tailed Wilcoxon rank-sum test, see Methods) component(s) to unpredicted reward delivery. Bottom: 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. **h** Top: Mean ΔF/F aligned to unpredicted reward delivery with quasi-simultaneous 405nm (black) and 470nm (blue) illumination in a mouse expressing the non-functional mutant ACh3.0 sensor. Shaded regions, S.E.M. Bottom: Histogram of Pearson's correlation coefficients between ΔF/F fluorescence traces obtained with quasi-simultaneous 405nm and 470nm illumination in mutant ACh sensor expressing mice (n = 3). **i** Mean unpredicted reward triggered ΔF/F from a single fiber in functional ACh3.0 sensor (top) and mutant ACh sensor (bottom) 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.

Fig. 2

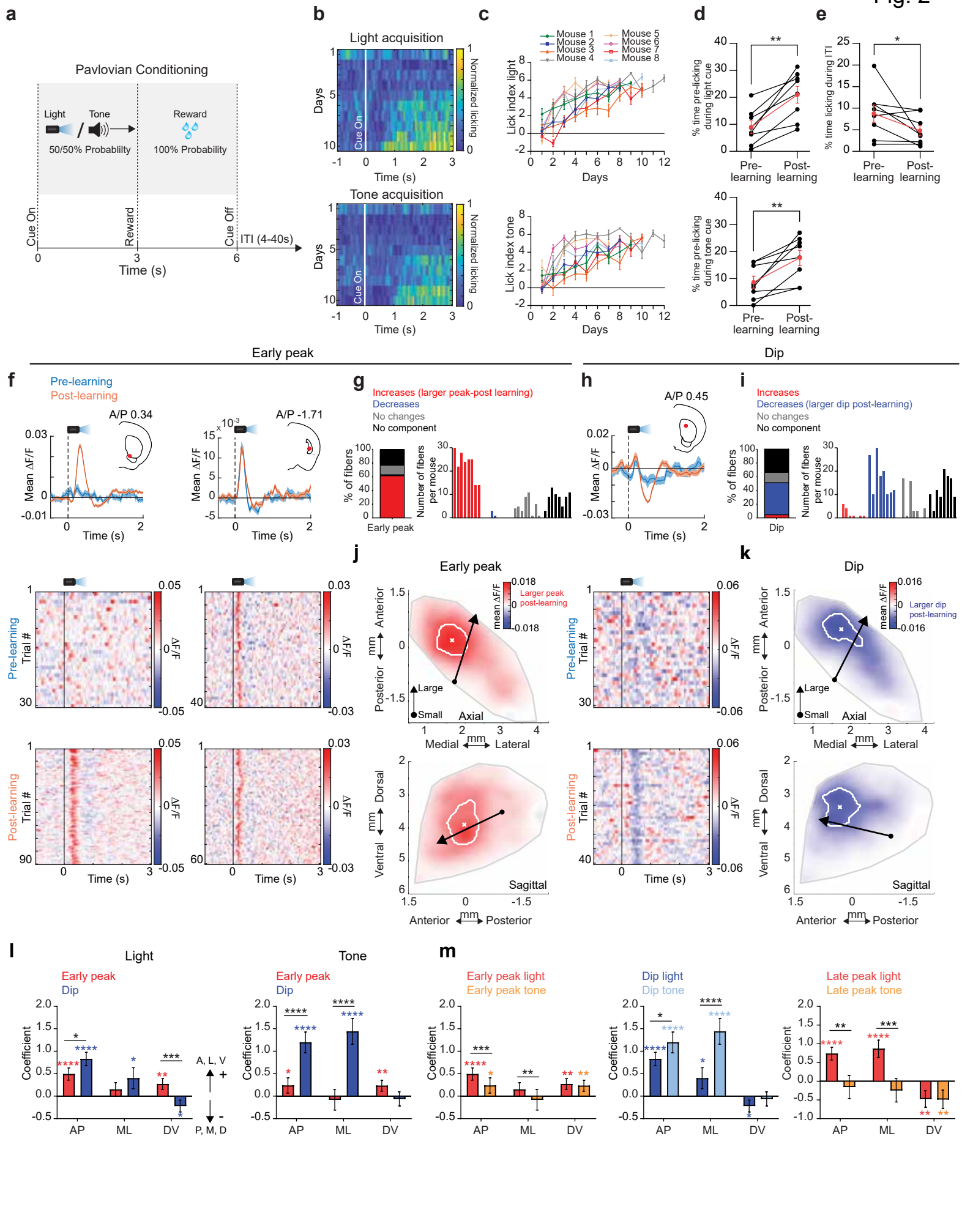


Fig. 2: ACh release components for conditioned cues evolve bi-directionally over distinct anatomical gradients during Pavlovian learning. a Schematic of dual-cue delay Pavlovian conditioning task. **b** Normalized mean lick spout contacts aligned to light (top) and tone (bottom) cue onset for 10 consecutive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

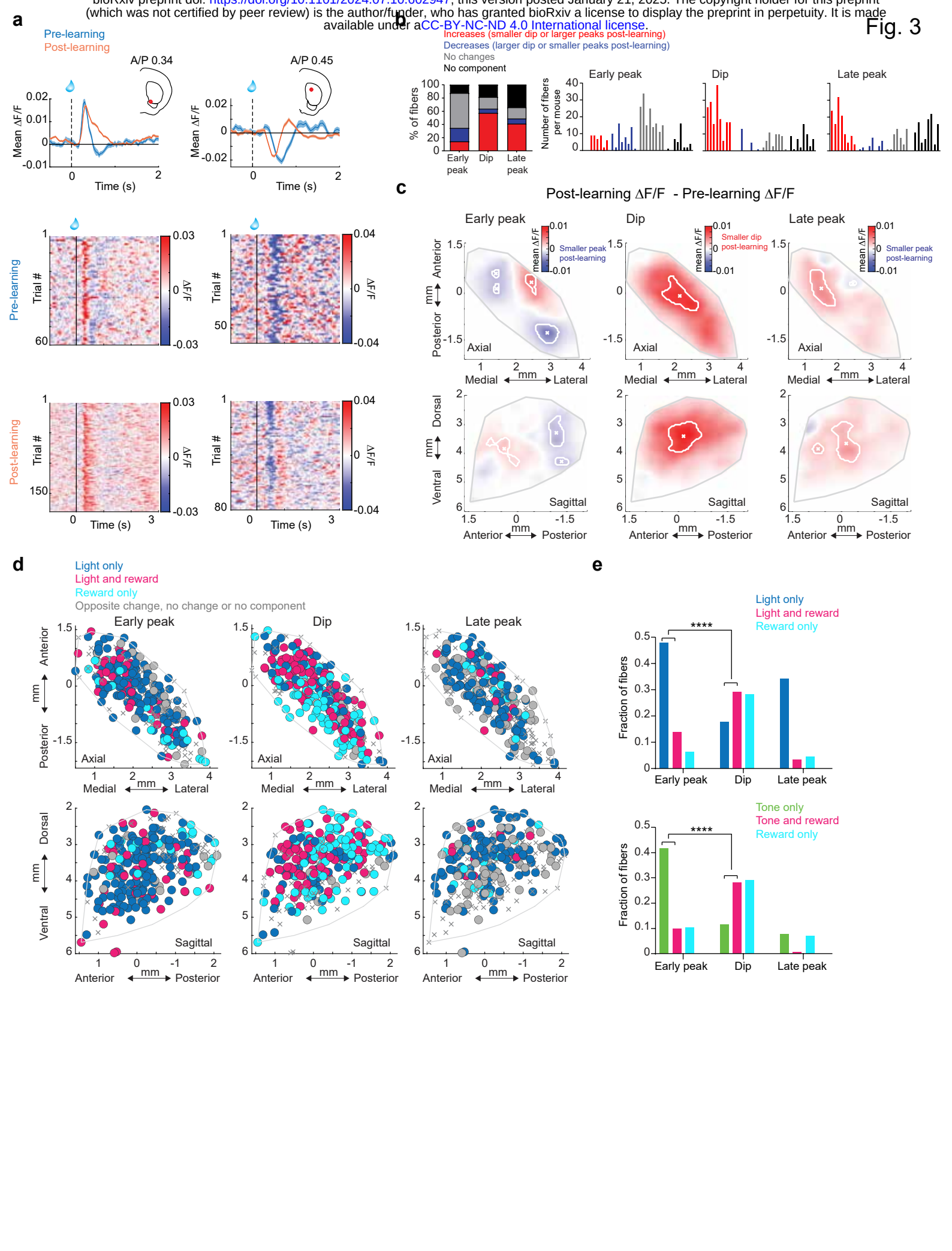


Fig. 3: Opposite ACh changes with learning at cue and reward are signal component selective and spatially concentrated in the anterior dorsal striatum. a Top: Mean ΔF/F aligned to the reward delivery in the task for 2 representative fibers (same fibers as Fig. 2f, left and 2h) for trials pre (blue) and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

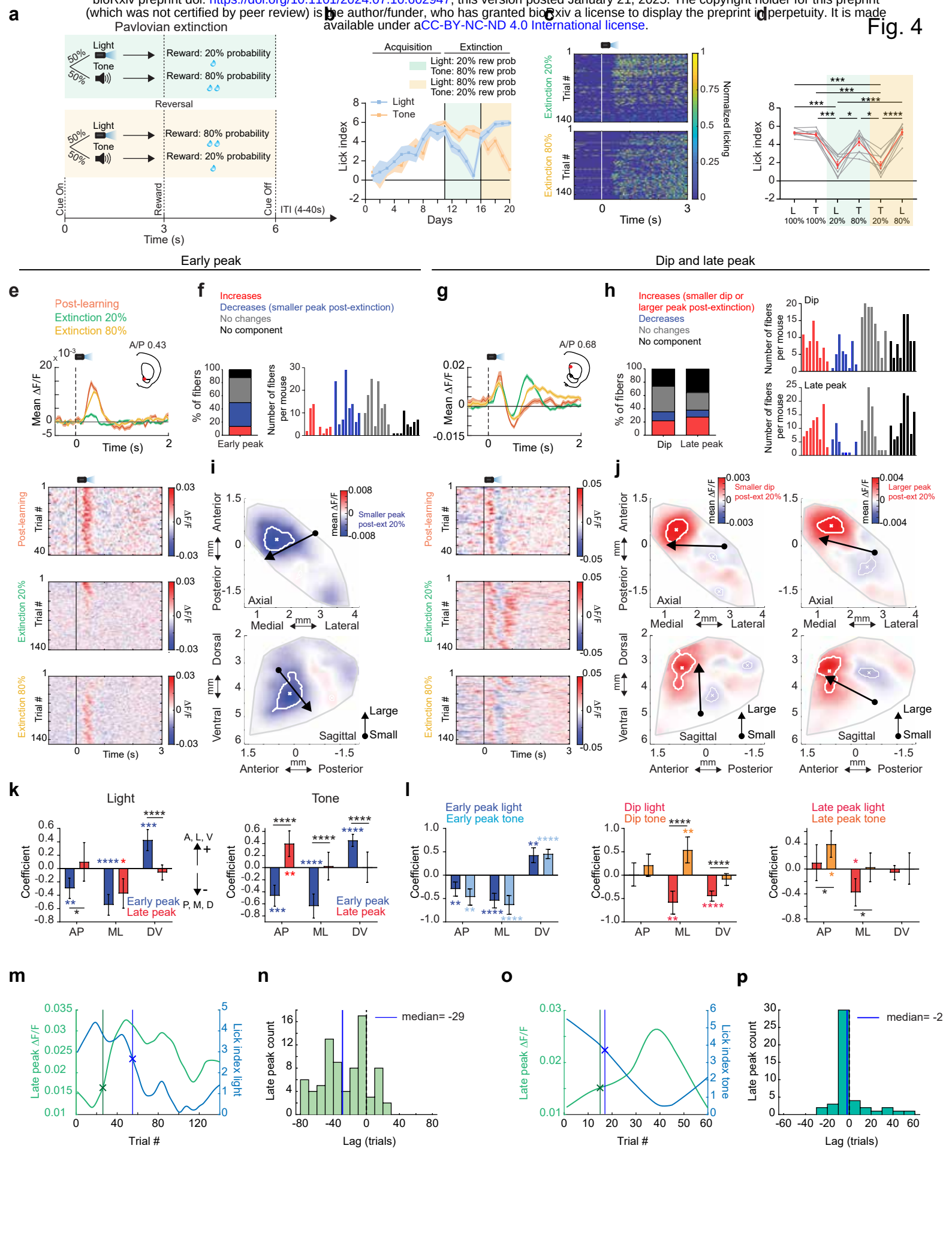


Figure 4: Emergence of cue-evoked ACh increases in the anterior dorsal striatum during partial extinction. a Schematic of the partial extinction paradigm. **b** Mean lick index on light and tone cue trials for each day of training for one representative mouse. Shaded regions, S.E.M. across trials for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

Fig. 5

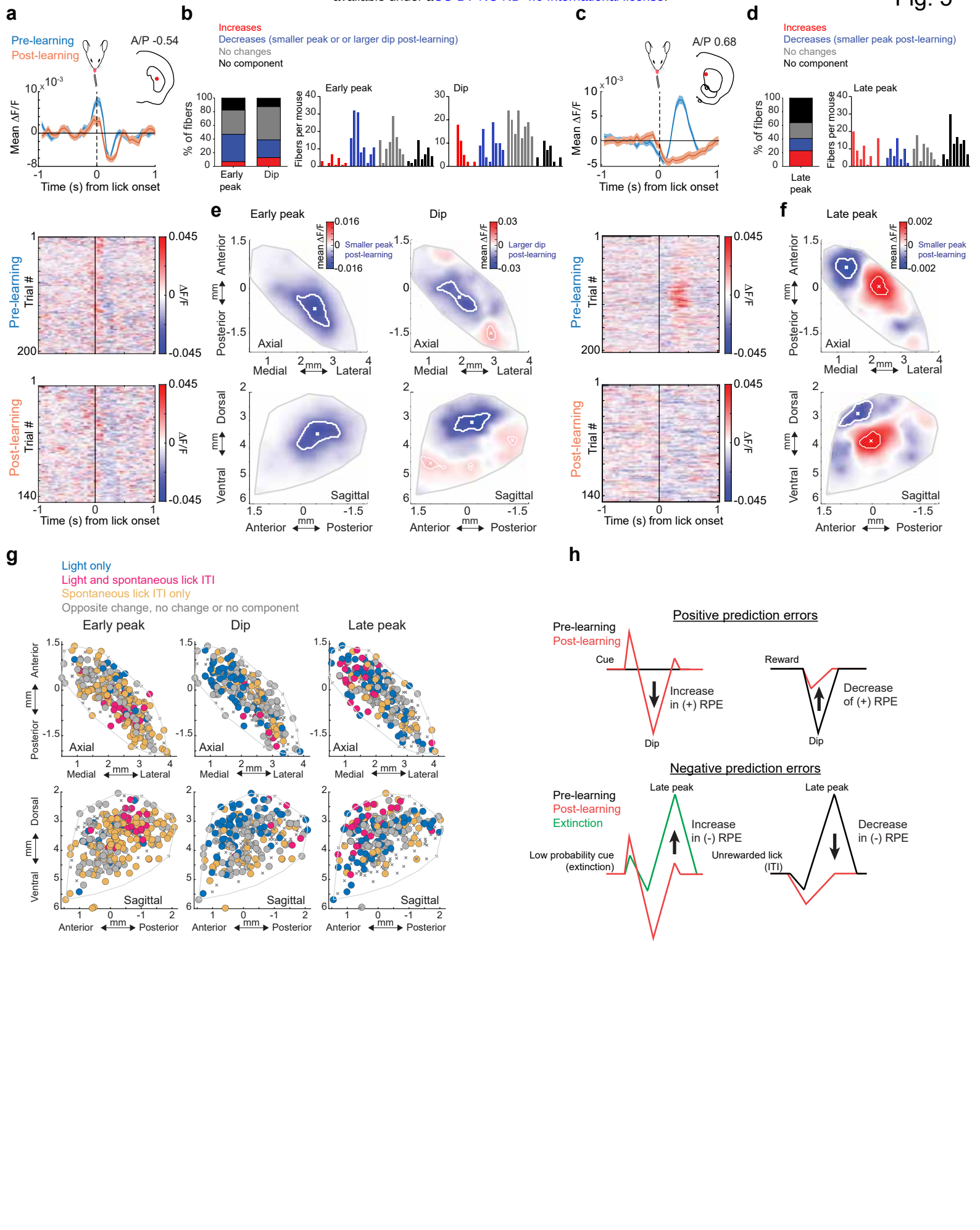


Fig. 5: Regional changes in ACh release components to unrewarded ITI licks during learning. a Top: Mean ΔF/F aligned to unrewarded ITI lick bout onsets for one representative fiber with significant early peak and dip changes for trials pre (blue) and post (orange) Pavlovian learning. Shaded regions,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

Fig. 6

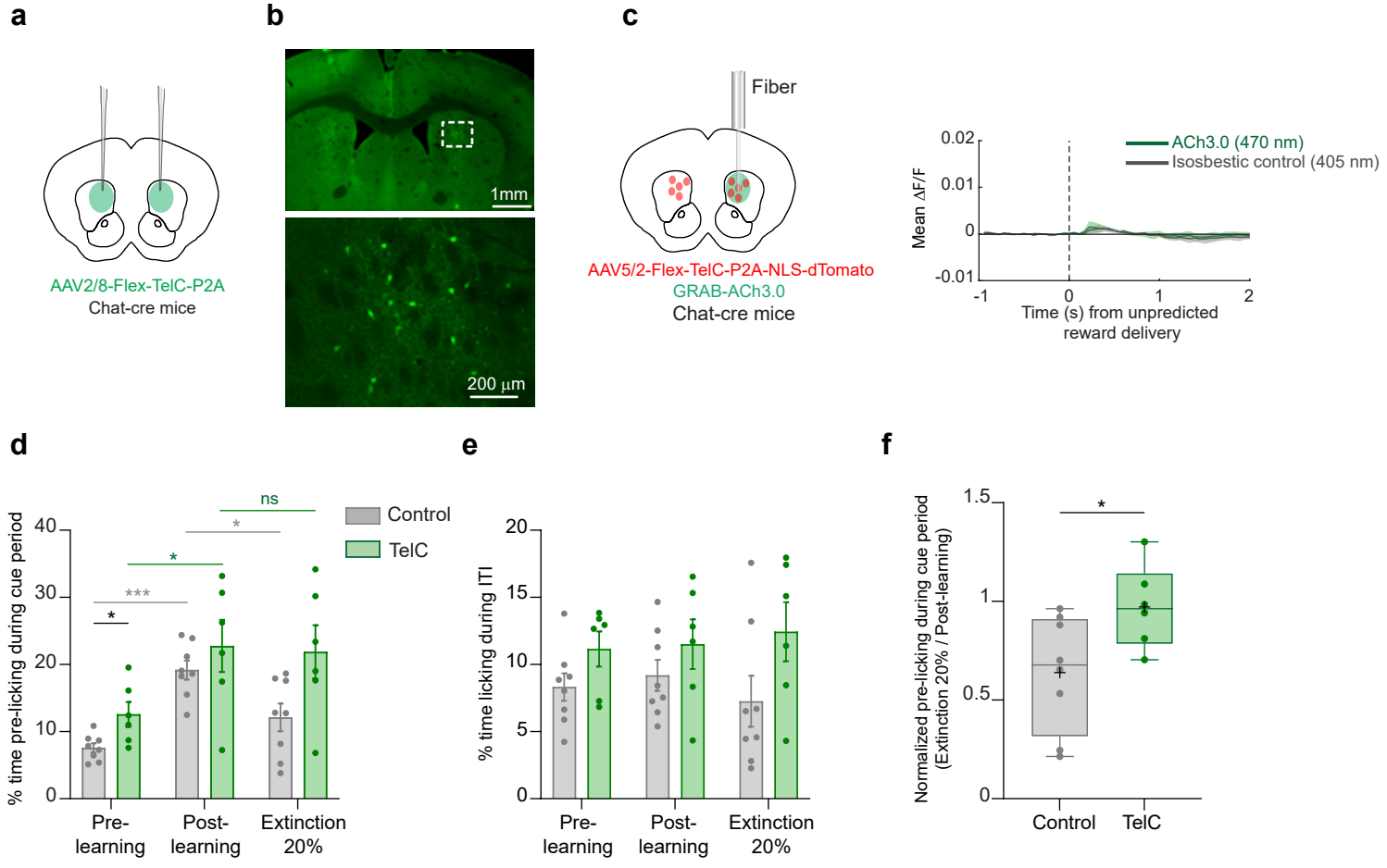


Fig. 6: **Inactivation of ACh release in the anterior dorsal striatum impairs behavioral changes during extinction. a** Schematic showing the bilateral injection strategy for the selective expression of TelC in CINs in aDS. **b** Left: Fluorescence image in a coronal section of the striatum showing bilateral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

Fig. 7

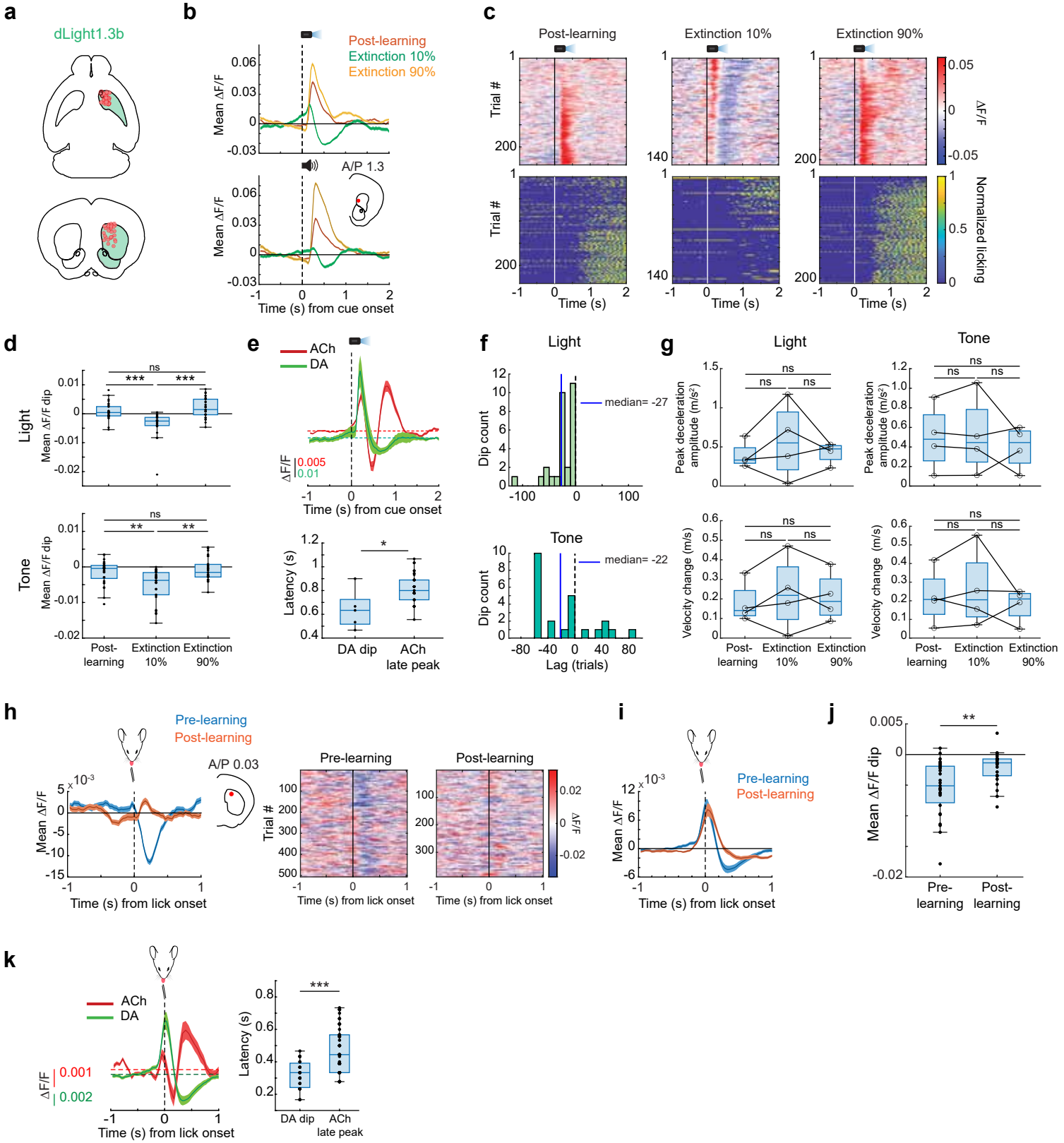


Fig. 7: Dips in dopamine release in the anterior dorsal striatum are present for unrewarded lick onsets before learning and for cues after extinction. a Schematic showing the locations of fibers (n = 29 fibers in 4 mice) in the aDS for measurements of DA release with dLight1.3b. **b** Light (top) and tone (bottom) cue onset triggered averages of ΔF/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 ΔF/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, Friedman test, $p > 0.05$. Each datapoint is the mean for one mouse. **h** Left: Mean ΔF/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 ΔF/F for all bout onsets included in triggered average on left. **i** Mean lick bout onset-triggered ΔF/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) ΔF/F in the aDS (see Methods) for bouts pre learning. Shaded region, S.E.M. Right: box plot of latencies to the minimum (DA dip) or maximum (ACh peak) for all fibers in DA or ACh sensor-expressing mice respectively in the pre learning phase. Each dot is one fiber. ***$p<$ 0.001, Two-tailed Wilcoxon rank-sum test.

Fig. 8

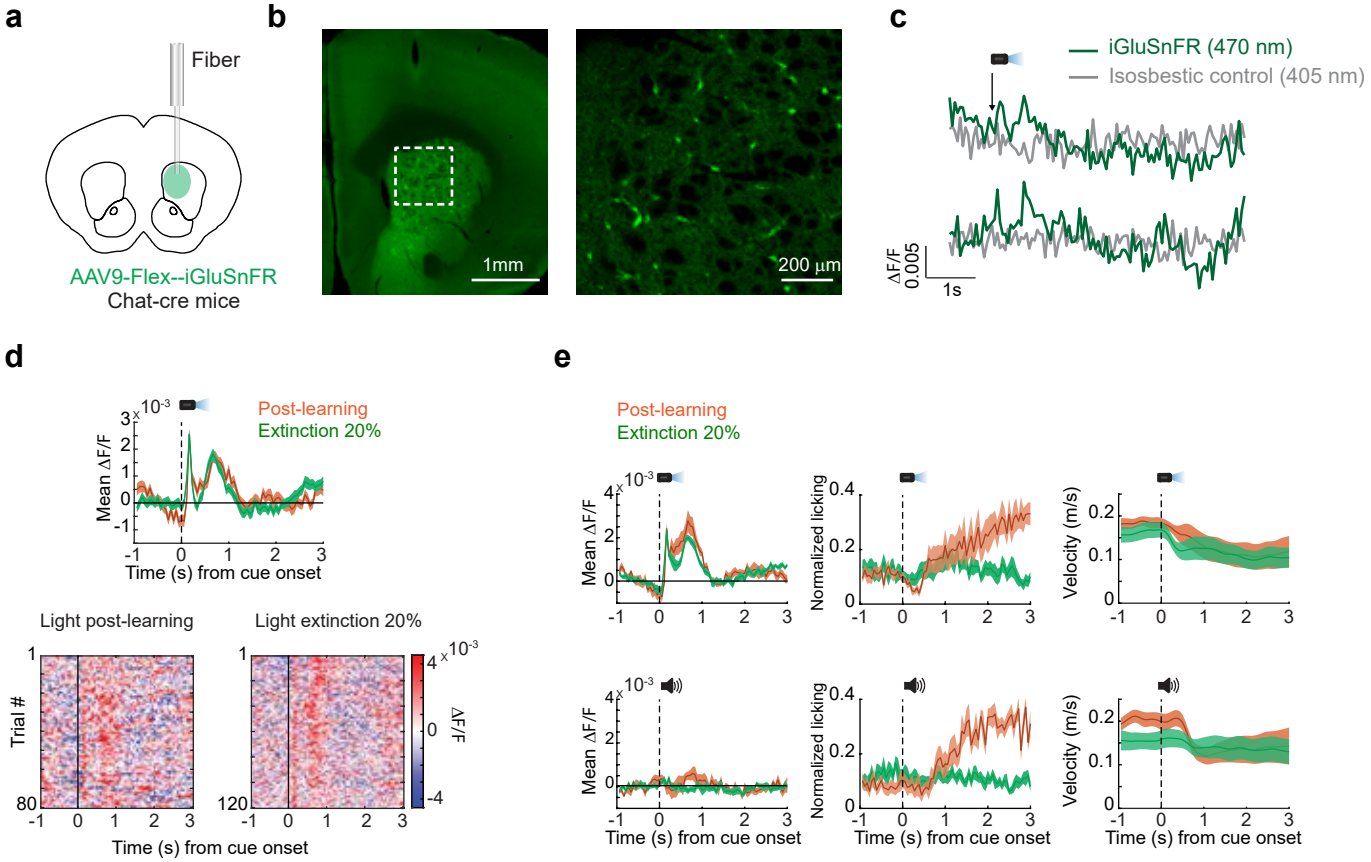
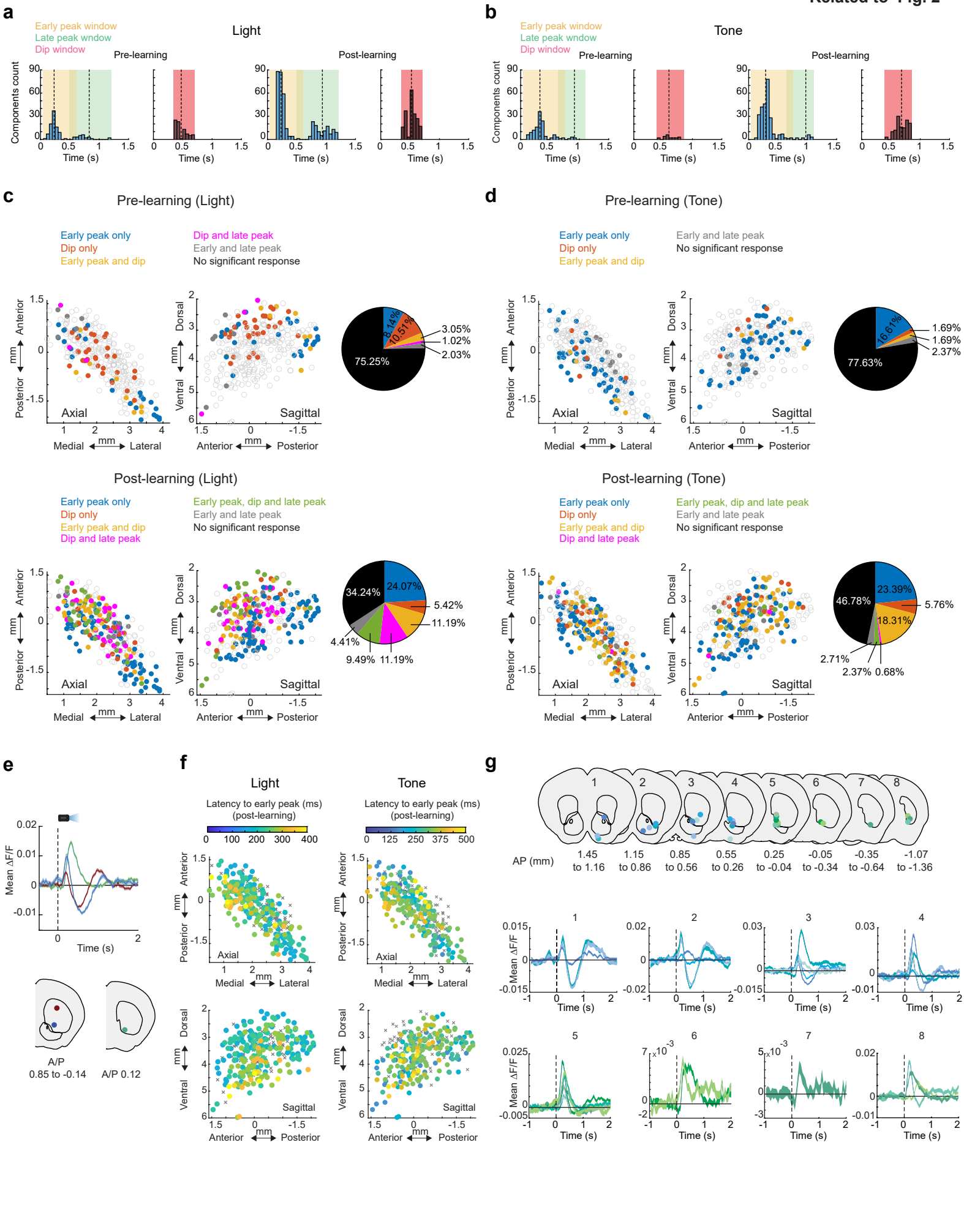
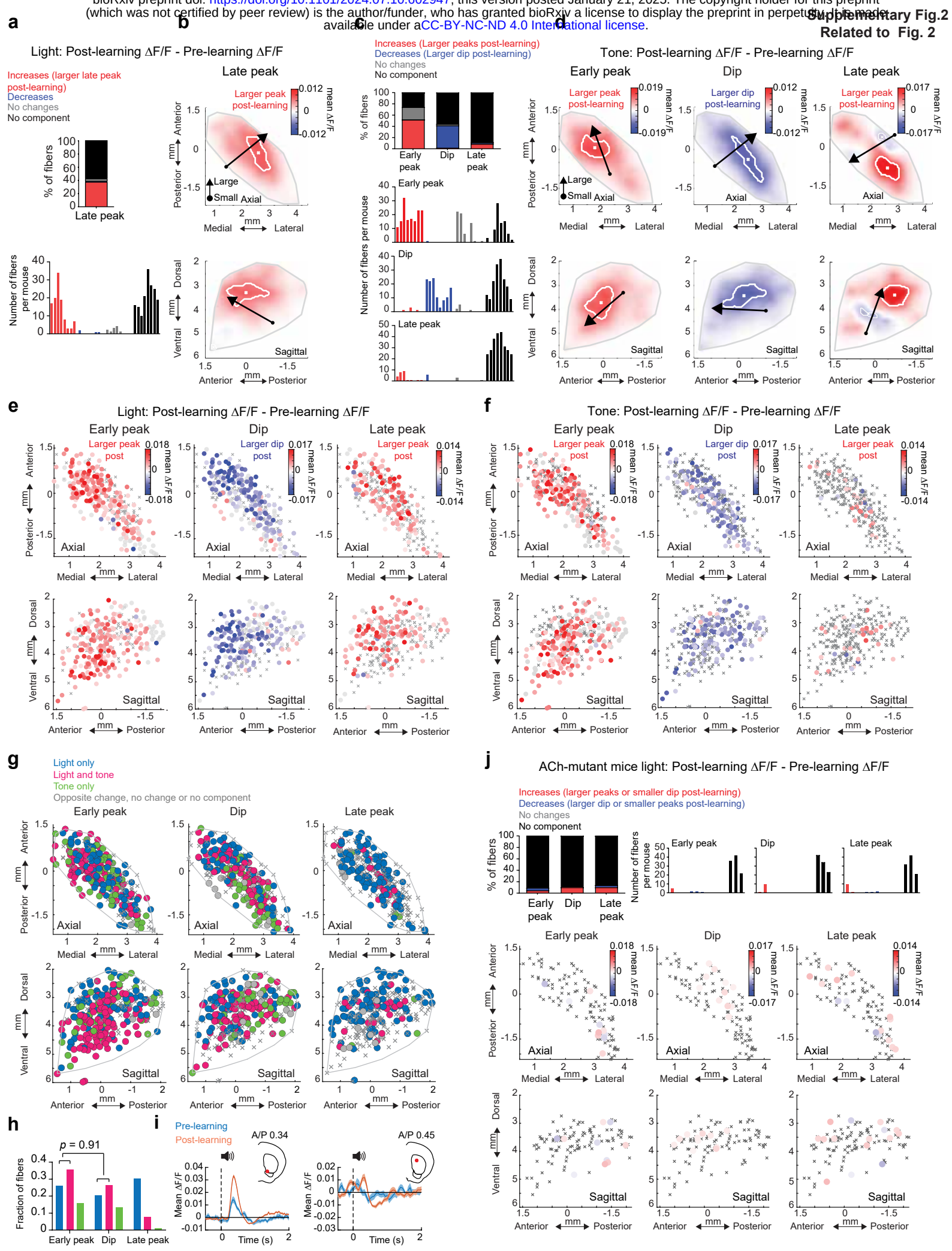


Fig. 8: Cue-evoked glutamate release onto anterior striatum cholinergic interneurons does not change during extinction. **a** Schematic showing strategy for single fiber optical measurements of bulk glutamate release selectively on aDS CINs with iGluSnFR in ChAT-cre mice. **b** Left: fluorescence image 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 ΔF/F traces aligned to light cue onset for two trials in one mouse. **d** Top: mean ΔF/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 ΔF/F for all trials included in the triggered averages at top. **e** Mean ΔF/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.

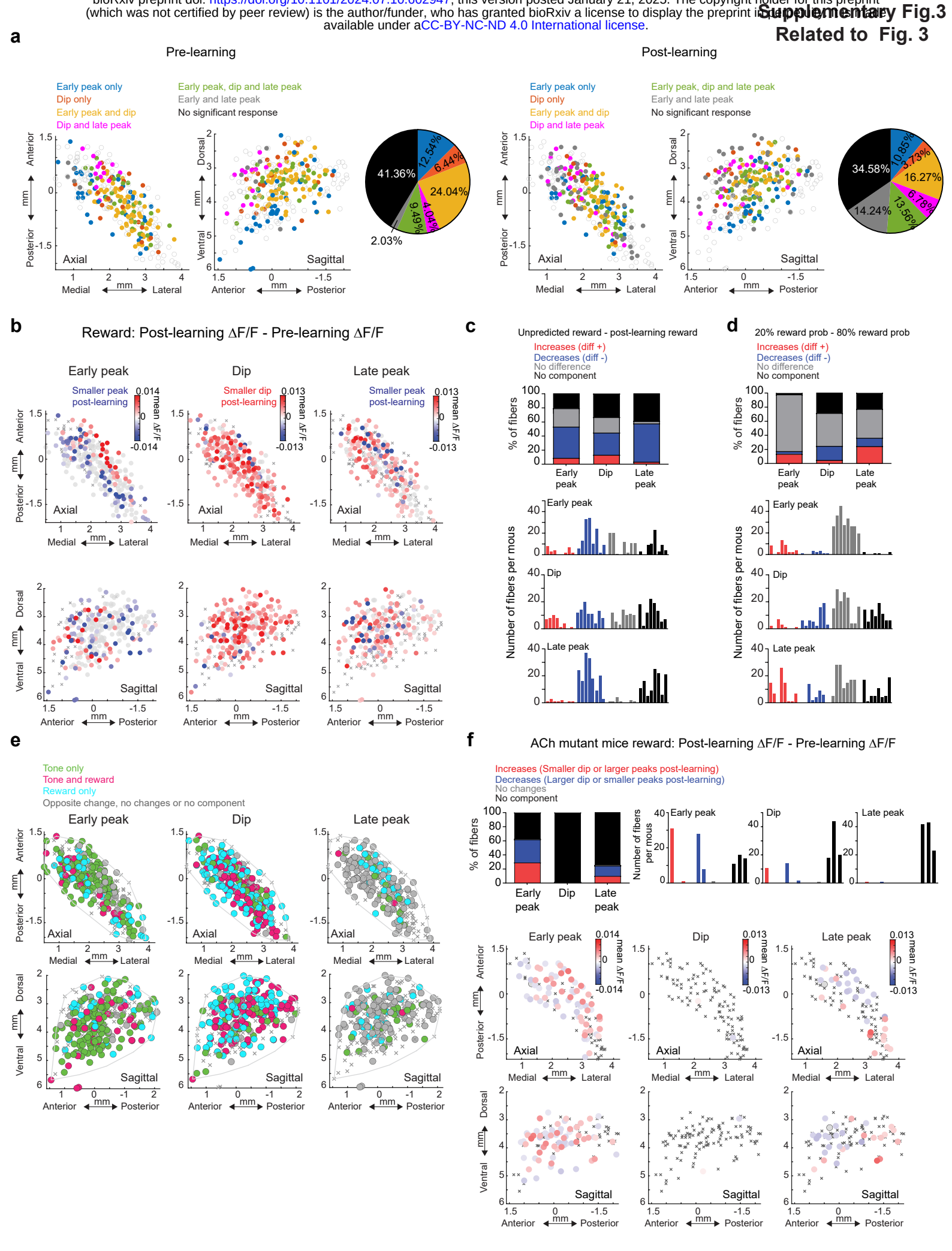


SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

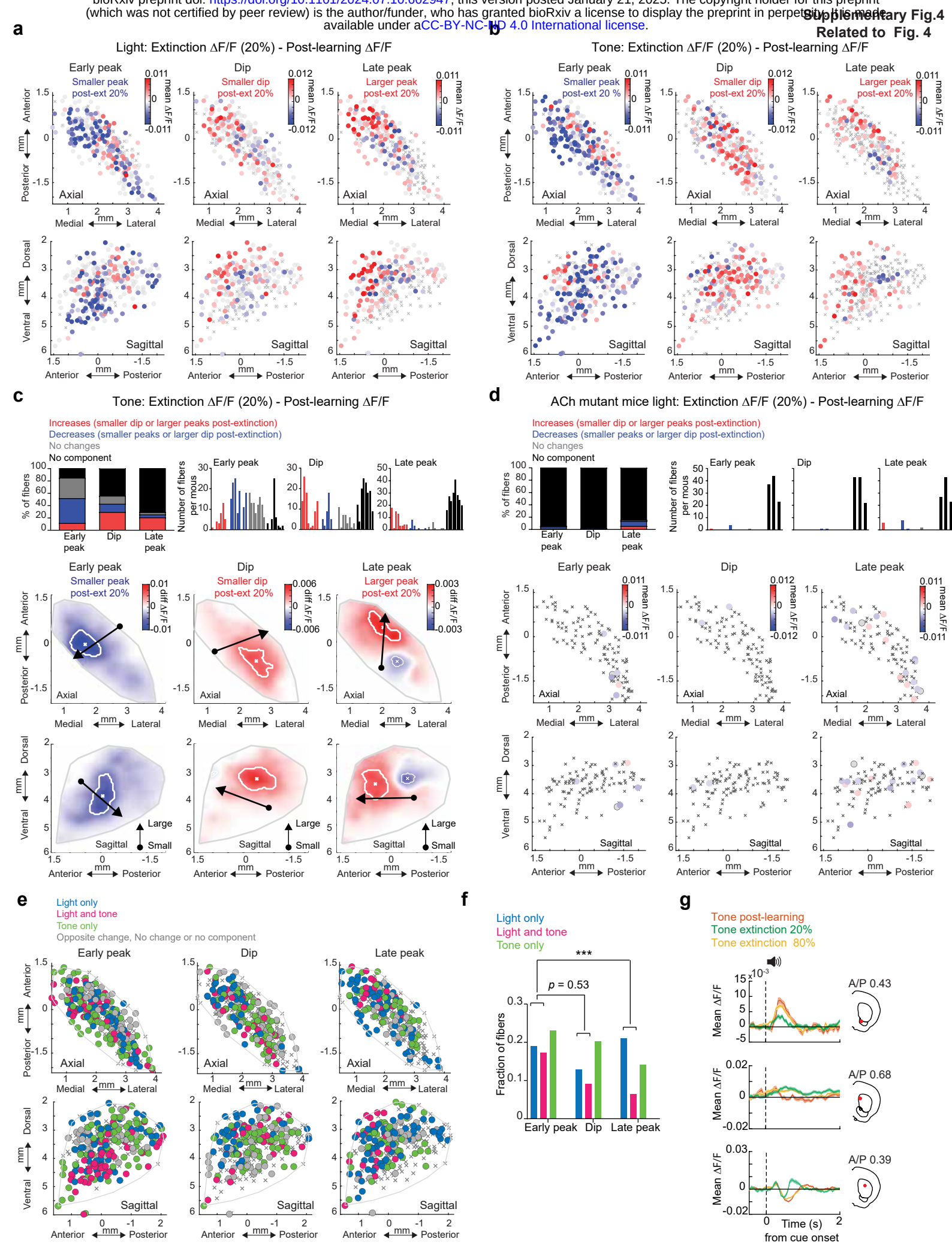
Supplementary Fig. 1: Additional quantification of component specific changes in cue-evoked ACh signals and mutant ACh-sensor controls during Pavlovian learning. a Histogram of latencies to significant (see Methods) peaks and troughs to light cue onset pre (left) and post (right) learning. Shaded regions indicate time windows used to identify each component (early peak, dip, and late peak), and the dashed lines indicate the means (pre learning: early peak 0.23± 0.008s, dip 0.47±0.009s, late peak 0.82±0.045s; post learning: early peak 0.23±0.003s, dip 0.52±0.005s, late peak 0.93±0.012s) **b** Same as (**a**) for the tone cue (pre learning: early peak 0.32±0.01s, dip 0.6±0.02s, late peak 0.91±0.02s; post learning: early peak 0.23±0.005s, dip 0.52±0.009s, late peak 0.93±0.03s). T-test for light vs. tone component latencies: pre learning early peak: $p = 9.3 \times 10^{-10}$, post learning early peak $p = 3.78 \times 10^{-19}$; pre learning dip $p = 2.07 \times 10^{-7}$, post learning dip $p = 1.97 \times 10^{-41}$; pre learning late peak $p = 0.47$, post learning late peak $p = 0.33$. **c** Maps (axial left, sagittal right) indicating the presence of each combination of signal components to light cue onset for each fiber (circle) pre (top) and post (bottom) learning. Empty circles are fibers with no significant response. Pie charts indicate percentages of fibers with each combination of signal components for light cue onset pre (top) post (bottom) learning**. d** Same as (**c**) for the tone cue. **e** Mean light cue triggered ΔF/F for post learning trials for three representative fibers in the same mouse (locations in coronal sections at bottom) illustrating the larger magnitude and longer latency ACh peaks in more ventral and posterior locations and relatively larger dips in more anterior and dorsal regions, consistent with the patterns shown in Fig. 2. **f** Maps (axial top, sagittal bottom) of latencies to early peaks to light (left) and tone (right) cue onsets post learning for all fibers. Xs are fibers with no significant early peak. **g** Light cue onset-triggered ΔF/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.



Supplementary Fig. 2: Additional quantification of the spatial organization of light and tone cue-evoked ACh release changes during learning. a Top: percent of all fibers with significant increases or decreases, no change, or no significant component from pre to post-learning for the late peak ΔF/F at light cue onset. Bottom: histogram of number of fibers per mouse with significant late peak changes. Each bar is the fiber count for one mouse for each condition indicated by colors at left. **b** Maps (axial, top; sagittal, bottom) showing spatially weighted means across locations of differences with learning (post - pre ΔF/F) for the mean late peak ΔF/F at light cue onset. Lines indicate the axes of maximal variation and arrows indicate the direction of peak increases from smallest to largest changes. White contours indicate regions with changes in the highest 10th percentile. **c** Top: percent of all fibers with significant ΔF/F increases or decreases, no change, or no significant component from pre to post learning for each component at tone cue onset. Bottom: histogram of number of fibers per mouse with significant changes for each component. Each bar is the fiber count for one mouse for each condition indicated by colors at top. **d** Maps (axial, top; sagittal, bottom) showing spatially weighted means across locations of differences with learning (post-pre ΔF/F) for the mean ΔF/F for the three signal components at tone cue onset. Lines indicate the axes of maximal variation and arrows indicate the direction of components increases from smallest to largest changes. White contours indicate regions with changes in the highest 10th percentile. **e** Maps showing the difference in mean light-cue evoked ΔF/F with learning (post – pre) for the three signal components for each fiber (circle). Gray circles, no significant difference; Xs, no significant component. Significance calculated with two-tailed Wilcoxon rank-sum test ($p < 0.01$). **f** Same as (**e**) for tone cue. **g** maps showing each fiber (dot) color coded according to whether significant changes from pre to post learning were present for the light cue onset only (dark blue), tone cue onset only (green) or both (pink). **h** Fraction of all fibers for each component classified according to changes across learning for light cue, tone cue, or both as indicated in (**g**) ($p = 0.91$, for early peak vs. dip, Fisher's exact test). **i** Mean ΔF/F for 2 representative fibers (same as the fibers shown in Fig. 2f (left) and Fig. 2h) aligned to 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. **j** Top left: percent of all fibers in mice expressing the non-functional mutant ACh sensor with significant ΔF/F increases or decreases, no change, or no significant component from pre to post learning for each component at light cue onset. Top right: histogram of number of fibers per mouse with significant early peak changes. Each bar is the fiber count for one mouse for each condition indicated by colors at left. Bottom: Maps showing the difference in mean light-cue evoked ΔF/F with learning (post – pre) for the three signal components for each fiber (circle) in mice expressing the mutant ACh sensor. Gray circles, no significant difference; Xs, no significant component. Significance calculated with two-tailed Wilcoxon rank-sum test ($p < 0.01$).

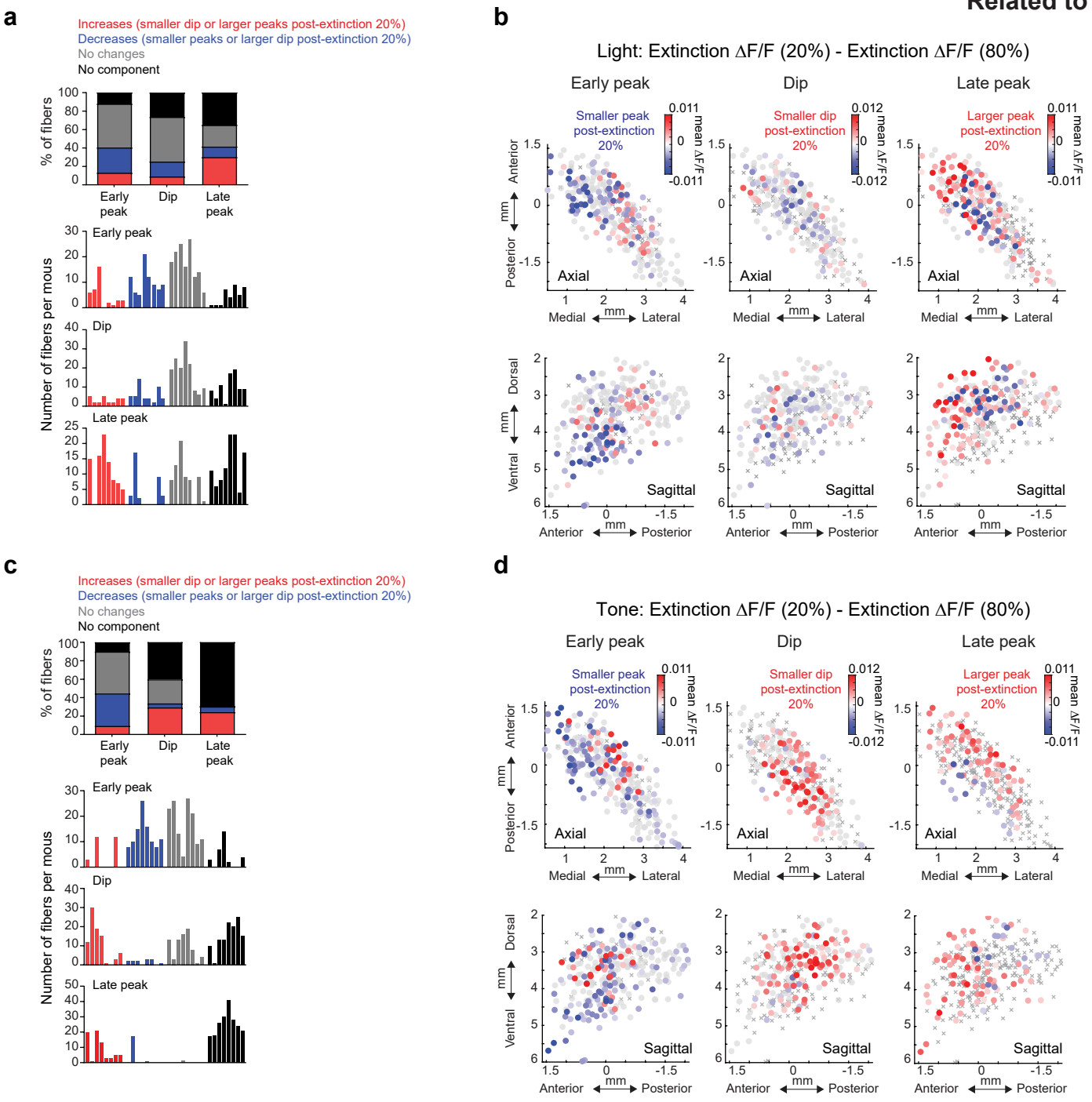


Supplementary Fig. 3: Additional quantification of changes in reward-evoked ACh release and mutant ACh-sensor controls during learning. a Maps indicating the presence of each combination of signal components to reward consumption onset for each fiber (circle) pre (left) and post (right) learning. Empty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



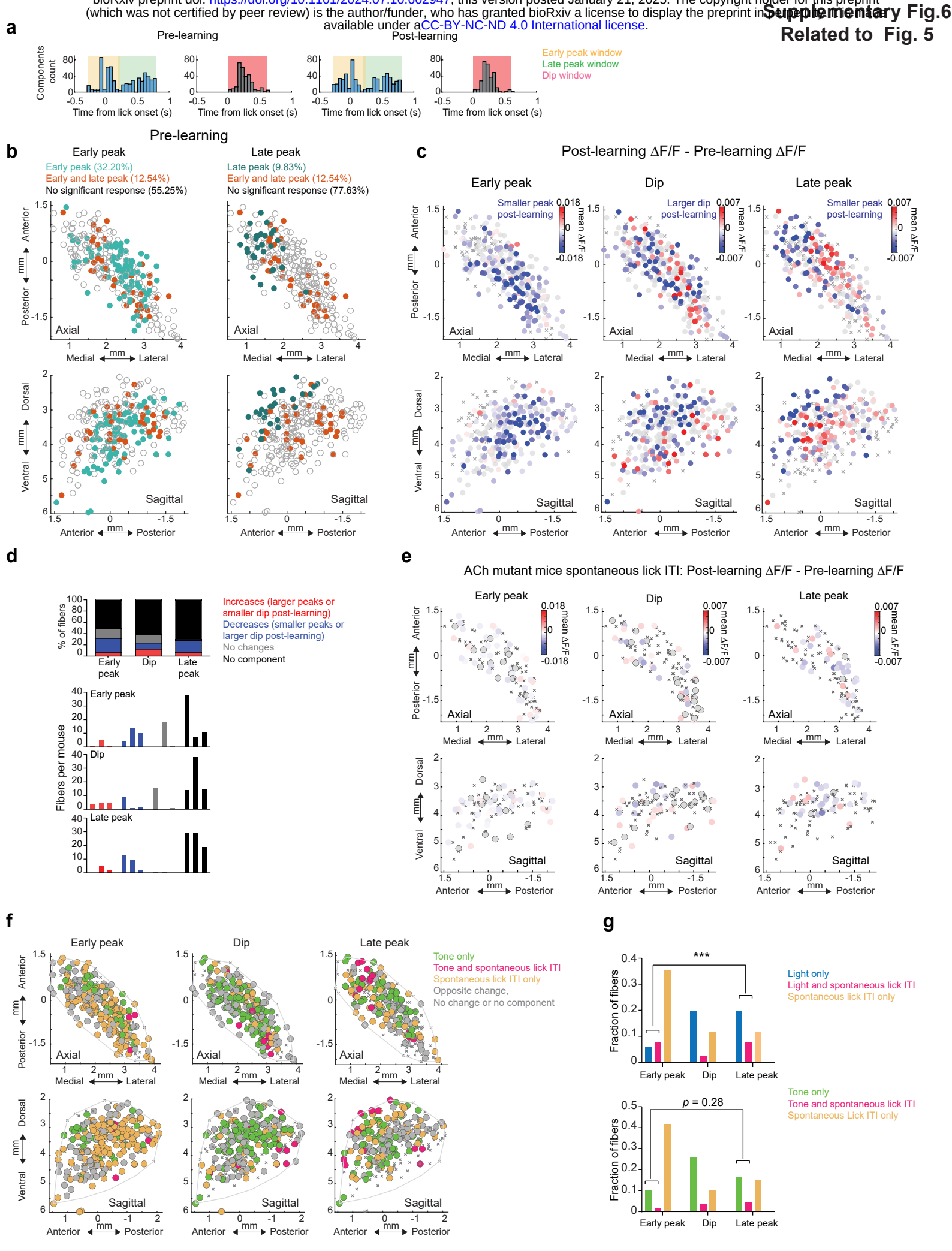
Supplementary Fig. 4: Additional quantification of changes in cue-evoked ACh release and mutant ACh-sensor controls during extinction. a Maps showing the difference in mean light cue-evoked ΔF/F between the post-learning (100%) and 20% reward probability phases for the three signal components. Each circle is one fiber. Gray circles, no significant difference; Xs, no significant component. Significance calculated with two-tailed Wilcoxon rank-sum test ($p < 0.01$). **b** Same as (**a**) for tone. **c** Top left: percent of all fibers with significant ΔF/F differences between the 20% extinction and 100% post learning phases (20% - 100% ΔF/F) at tone cue onset for each component. Top right: histogram of number of fibers per mouse with significant changes for each component. Each bar is the fiber count for one mouse for each condition indicated by colors at top. Bottom: Maps (axial, top; sagittal, bottom) showing spatially weighted means across locations of differences between the 20% extinction and 100% post learning phases (20% - 100% ΔF/F) for the mean early peak ΔF/F at tone cue onset. Lines indicate the axes of maximal variation and arrows indicate the direction of peak decreases from smallest to largest changes. White contours indicate regions with changes in the highest 10th percentile. **d** Top: fiber percentages and counts per mouse as in (**c**) for the light cue in mice expressing the non-functional mutant ACh sensor. Bottom: maps as in (**a**) for the mutant ACh sensor mice for the light cue. **e** Maps showing each fiber (circle) color coded according to whether significant differences between the 20% extinction and 100% post learning phases were present for the light cue onset only (dark blue), tone cue onset only (green) or both (pink). **f** Fraction of all fibers for each component classified according to changes across extinction for light cue, tone cue, or both as indicated at top ($p = 0.53$ for early peak vs. dip and $p = 0.0008$ for early peak vs. late peak; Fisher's exact test, *** $p<0.001$). **g** Mean ΔF/F for 3 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. Top and middle examples are the same fibers shown in Fig. 4e and g, respectively).

Supplementary Fig.5
Related to Fig. 4



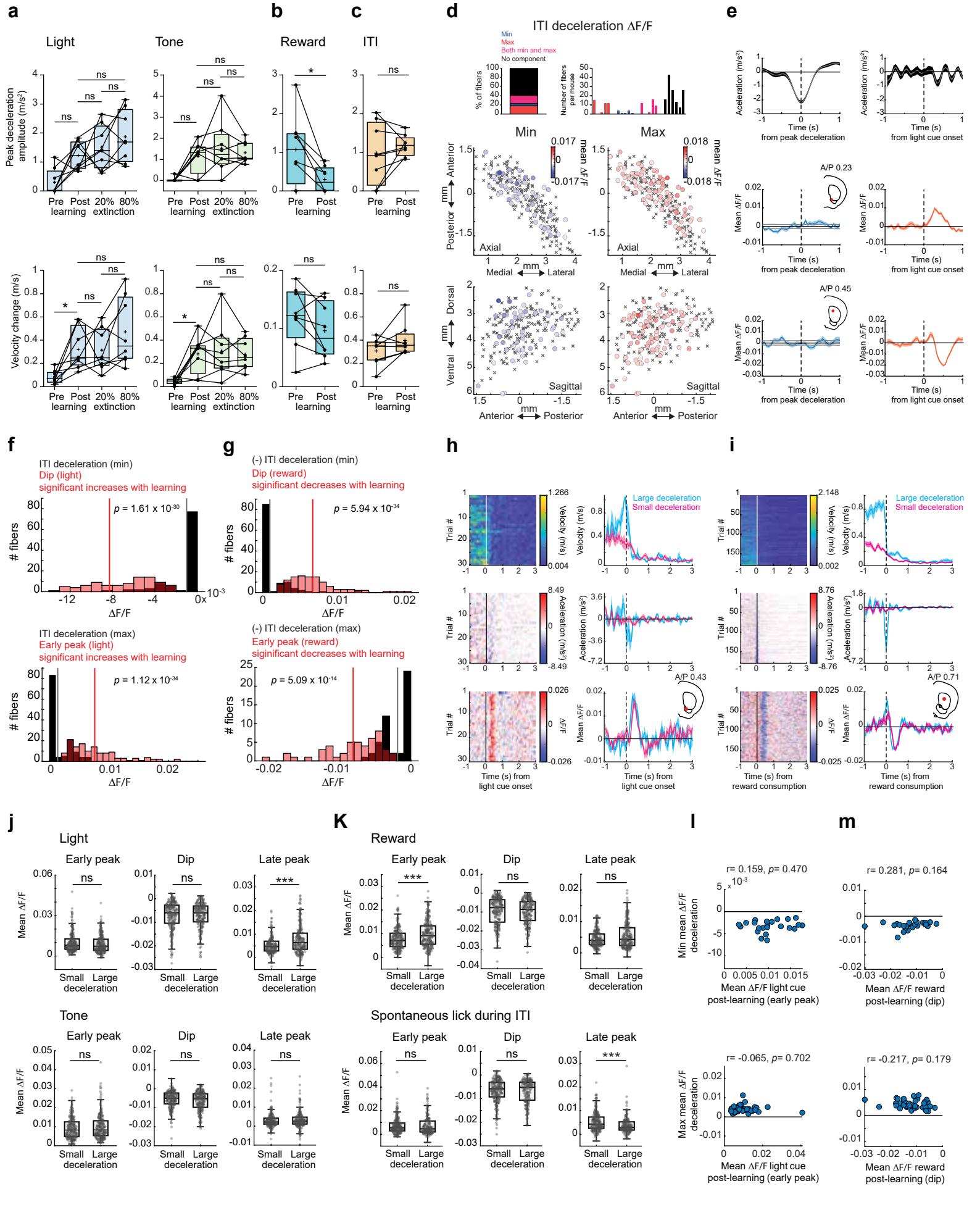
Supplementary Fig. 5: Comparison of ACh release for high and low probability cues during extinction sessions. a Top: percent of all fibers with significant ΔF/F differences between the 20% extinction and 80% extinction phases (20% - 80% ΔF/F) at light cue onset for each component. Bottom: histogram of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

Supplementary Fig.6
Related to Fig. 5



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

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



Supplementary Fig. 7: Learning and extinction related changes in striatum-wide ACh release cannot be explained by locomotion changes. **a** Box and whisker plots showing average peak decelerations (top) and velocity changes (bottom) across all mice (n=8 mice) following light (left) tone (right) cue onsets for each learning phase. (n.s., not significant Friedman test, p >0.05, p = 0.04 for light and tone velocity, *p<0.05). Each datapoint is the mean for one mouse. **b** Same as (**a**) for reward delivery pre and post learning (Two-tailed Wilcoxon rank-sum test, p = 0.03 for peak deceleration and p = 0.14 for velocity; *p<0.05). **c** Same as (**b**) for unrewarded lick bout onsets in the ITI (Two-tailed Wilcoxon test, p = 0.38 for peak deceleration and p = 0.19 for velocity). **d** Top left: Percent of all fibers with significant minima, maxima, or both (min and max) of the mean ACh ΔF/F within a ±1-s window triggered on troughs of large spontaneous decelerations in the intertrial interval (ITI). Top right: Histogram of the number of fibers per mouse with significant changes for minima, maxima, or both. Bottom: Maps (axial, top; sagittal, bottom) showing the minima (left) and maxima (right) of the mean ACh ΔF/F triggered on troughs of large ITI decelerations. Significance for each fiber (circle) was determined by comparison to the 95% confidence interval of a bootstrapped ΔF/F distribution (Methods). Xs, no significant response. **e** Left column: Mean acceleration (top) and mean ΔF/F (middle and bottom) aligned to ITI deceleration troughs for 2 representative fibers with small but significant minima and maxima. Shaded regions, S.E.M. Gray shading indicates the 95 % confidence interval. Insets indicate fiber location in the coronal plane. Right column: Mean acceleration (top) and mean ΔF/F (middle and bottom) aligned to light cue onset for trials post-learning for the same representative fibers shown in the left column. Shaded region, S.E.M. **f** Red: Histogram showing the magnitudes of ΔF/F changes from pre to post learning for the dip (top) and early peak (bottom) components for the light cue for fibers with significant changes. Black: Histogram of the magnitudes of the ΔF/F minima (top) and maxima (bottom) of the spontaneous ITI deceleration triggered ΔF/F for the same fibers with changes over learning plotted in red. Fibers with no significant deceleration ITI response counted in the 0 bin. Lines indicate means (Two-tailed Wilcoxon rank-sum test). **g** Same as (**f**) for reward delivery periods. The sign of the ITI deceleration response magnitudes is flipped for visualization because the deceleration magnitudes to reward delivery decrease with learning, opposite to cues. **h** Left column: Total treadmill velocity (top), acceleration (middle), and ΔF/F (bottom) aligned to light cue onsets for all trials post-learning for a single example fiber with a significant early peak increase with learning sorted by the deceleration magnitude following cue onset. Right column: Mean cue-triggered averages for measures in left column for trials with large decelerations (>80th percentile of decelerations from the distribution of decelerations across all mice, n = 4 trials, blue) and small decelerations (<20th percentile, n = 8 trials, magenta). Shaded region, S.E.M**. i** Same as (**j**) but aligned on reward consumption for another example fiber with a significant change in the reward dip magnitude with learning. (n = 24 and 31 trials for large and small deceleration trials respectively). **j** Box and whisker plots showing the mean light (top) and tone (bottom) cue-evoked ΔF/F for each component (early peak, dip and late peak) for trials with small and large decelerations across all fibers (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** same as (**j**) for reward (top) and spontaneous ITI licks (bottom). **l** Scatterplots of the minimum (top) and maximum (bottom) of the mean ΔF/F triggered on ITI deceleration troughs and the mean post-learning light cue-evoked early peak ΔF/F for all fibers (dots) with significant components for both cues and ITI decelerations. R and p values are from Pearson's correlations. **m** same as (**l**) but for the dip component during reward post-learning.

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

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

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