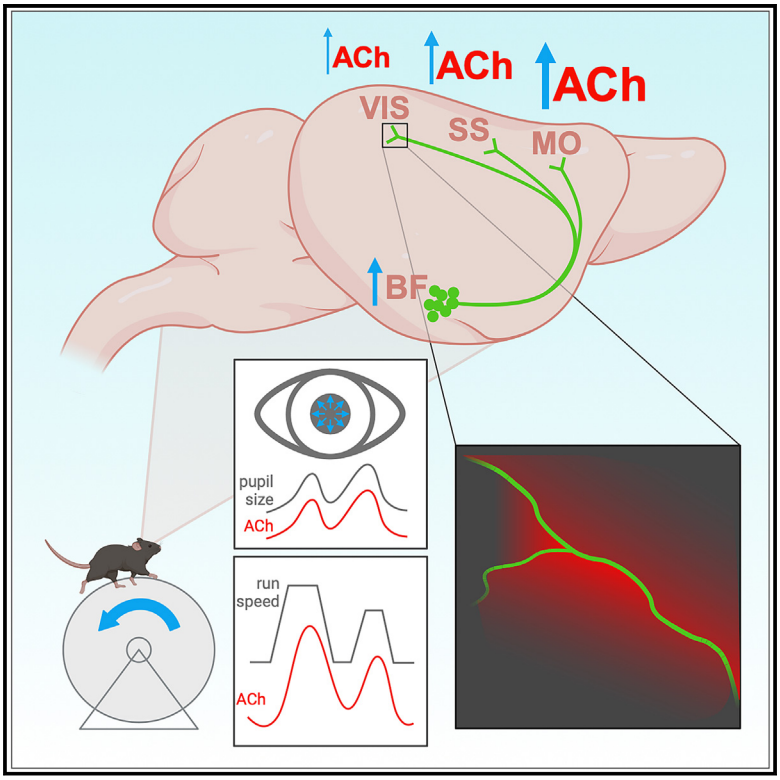
Cell Reports

Article

Cortical acetylcholine dynamics are predicted by cholinergic axon activity and behavior state

Graphical abstract



Authors

Erin Neyhart, Na Zhou, Brandon R. Munn, ..., Matthew J. McGinley, James M. Shine, Jacob Reimer

Correspondence

reimer@bcm.edu

In brief

Neyhart et al. describe precise timing and spatial characteristics of cortical acetylcholine availability in relation to cholinergic axon activity and behavioral markers.

Highlights

- Bulk ACh release around cortical cholinergic axons can be observed *in vivo*
- ACh levels scale with locomotion speed and pupil size
- Sensor deconvolution allows evaluation of precise ACh temporal dynamics
- ACh activity can be predicted from axon activity and behavioral variables

Neyhart et al., 2024, Cell Reports *43*, 114808
October 22, 2024
https://doi.org/10.1016/j.celrep.2024.114808

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Cortical acetylcholine dynamics are predicted by cholinergic axon activity and behavior state

Erin Neyhart,[1] Na Zhou,[1] Brandon R. Munn,[2,3] Robert G. Law,[1] Cameron Smith,[1] Zakir H. Mridha,[1] Francisco A. Blanco,[1] Guochuan Li,[4,5,6] Yulong Li,[4,5,6] Ming Hu,[1] Matthew J. McGinley,[1] James M. Shine,[2,3] and Jacob Reimer[1,7,*]

[1]Neuroscience Department, Baylor College of Medicine, Houston, TX 77030, USA
[2]Brain and Mind Centre, School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW 2006, Australia
[3]Complex Systems Group, School of Physics, Faculty of Science, The University of Sydney, Sydney, NSW 2006, Australia
[4]State Key Laboratory of Membrane Biology, Peking University School of Life Sciences, Beijing 100871, China
[5]PKU–IDG/McGovern Institute for Brain Research, Beijing 100871, China
[6]Peking–Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[7]Lead contact
*Correspondence: reimer@bcm.edu
https://doi.org/10.1016/j.celrep.2024.114808

SUMMARY

Acetylcholine (ACh) is thought to play a role in driving the rapid, spontaneous brain-state transitions that occur during wakefulness; however, the spatiotemporal properties of cortical ACh activity during these state changes are still unclear. We perform simultaneous imaging of GRAB-ACh sensors, GCaMP-expressing basal forebrain axons, and behavior to address this question. We observed a high correlation between axon and GRAB-ACh activity around periods of locomotion and pupil dilation. GRAB-ACh fluorescence could be accurately predicted from axonal activity alone, and local ACh activity decreased at farther distances from an axon. Deconvolution of GRAB-ACh traces allowed us to account for sensor kinetics and emphasized rapid clearance of small ACh transients. We trained a model to predict ACh from pupil size and running speed, which generalized well to unseen data. These results contribute to a growing understanding of the precise timing and spatial characteristics of cortical ACh during fast brain-state transitions.

INTRODUCTION

While animals are awake, their level of alertness and engagement with the world varies from moment to moment, and these spontaneous state changes can occur on fast timescales that are correlated with behavior and rapidly shift neural activity from one mode to another. For example, in the mouse visual cortex, neurons are depolarized and membrane potential variability is reduced during exploratory periods of locomotion and whisking,[1–4] visual responsiveness is enhanced,[1,3–7] and population noise and signal correlation are reduced.[8,9] Outside of these movement periods (i.e., during "quiet wakefulness"), there are also smaller changes in state with similar neural correlates, including a depolarized and less variable membrane potential, reduced amplitude of low-frequency membrane potential oscillations, and increased reliability and selectivity of evoked responses.[1,2] These covert state changes are tracked by dilation and constriction of the pupil that occur in the absence of any external changes in luminance.

Over the past decade, the mouse has served as an important model for understanding the mechanisms underlying these changes in state, which include neuromodulators[10] such as acetylcholine (ACh).[11–14] Indeed, activity in cholinergic axons from the basal forebrain (BF) projecting to the cortex is increased during whisking,[15–18] locomotion,[15,16,19,20] and pupil dilation outside of exploratory periods,[19] and genetically encoded ACh sensors have confirmed some of these effects. Axonal GCaMP activity might not have a straightforward relationship to ACh levels, which depend on local release probabilities and clearance kinetics as well as the kinetics of the fluorescent sensors themselves. Thus, more work is required to determine whether these measurements are equivalent and to precisely identify when and where ACh is available to influence the cortex during spontaneous fast state changes. Furthermore, the relationship with pupil and locomotion suggests that it should be possible to build a model that uses these easily measured variables to predict a significant amount of variance in cortical ACh. If this model was shown to generalize well, then it could then be employed in other head-fixed mouse experiments to estimate ACh fluctuations in the absence of ground-truth data.

We took advantage of recent advances in genetically encoded fluorescent sensor technology[21] to address these questions. We performed *in vivo* benchmarking of GRAB-ACh3.0 with exogenous application of ACh and characterized the relationship with spontaneous running speed and pupil size fluctuations. We performed simultaneous measurements of cholinergic axons, ACh activity, and behavior *in vivo*. Consistent with expectations that cholinergic axons are indeed releasing ACh, we found that ACh and axonal activity are tightly correlated and that, on a fine spatial scale, larger ACh increases can be

observed closer to cholinergic axons compared to farther away. We demonstrate an approach for deconvolving GRAB-ACh fluorescence to estimate the underlying changes in ACh levels—an approach which we believe could also be usefully applied to other fluorescent sensors that provide measurements of "analog" signals, like neuromodulator or neuropeptide levels. Finally, we show that ACh levels can be predicted from pupil size and treadmill speed with relatively high accuracy by a trained neural network.

RESULTS

Experimental setup and *in vivo* sensor validation

We used two-photon microscopy to measure ACh activity in the cortex of wild-type mice by implanting a cranial window over the visual cortex (VIS) and expressing the genetically encoded fluorescent ACh sensor GRAB-ACh3.0 via virus injection (Figure 1A). Although GRAB-ACh3.0 has been benchmarked against varying concentrations of ACh *in vitro*,[21] this benchmarking has not been performed *in vivo*. To measure the sensor response *in vivo*, we utilized an approach similar to two-photon-guided patching,[1] with a pipette entering through a hole in the coverslip. This technique allowed us to image GRAB-ACh signals in an anesthetized mouse while directly puffing small amounts of ACh into the cortex in the field of view to measure ACh sensor responses (Figure 1B; Video S1). We observed strong increases in GRAB-ACh fluorescence with 10 μM, 100 μM, and 1 mM ACh concentrations (Figure 1C). GRAB-ACh ΔF/F responses appeared to be larger with puffs of 100 μM ACh than 1 mM ACh, but this might be due to a saturating response of the sensor at 1 mM and the higher ACh concentration not fully washing out before the next puff was administered. Puffs of ACh-free cortex buffer and 100 μM norepinephrine resulted in mild decreases in GRAB-ACh fluorescence, presumably due to the transient wash-out of endogenous levels of ACh. Interestingly, puffs of 1 μM ACh also resulted in a decrease in fluorescence, suggesting that that baseline concentrations of ACh under isoflurane anesthesia in the VIS may be in the single millimolar range, consistent with microdialysis studies,[22] but higher than in some reports with cyclic voltammetry.[23] Overall, these experiments directly confirmed that changing ACh levels *in vivo* are rapidly tracked by the GRAB-ACh sensor.

We next repeated the measurements we had performed previously, relating cholinergic axons to locomotion and pupil under spontaneous conditions, but now using the GRAB-ACh sensor. The mouse was head fixed on a treadmill and allowed to run freely while the pupil was recorded with a camera, and we simultaneously imaged GRAB-ACh in the VIS (Figure 1D). As expected based on previous studies,[15,16,19,24] we observed strong increases in ACh levels that tracked spontaneous locomotion periods and the accompanying large pupil dilations (Figure 1E). Focusing on the small dilations that occurred outside of locomotion (i.e., during quiet wakefulness), we observed a high correlation between spontaneous ACh fluctuations and small dilations and constrictions of the pupil (Figure 1F).

This raised the question of whether we could precisely define the temporal relationship between changes in ACh levels and cholinergic axon activity by measuring them simultaneously. While neuromodulator axon activity and release have been measured simultaneously *in vitro*,[25] it is possible that *in vivo* axon-neuromodulator dynamics could differ. We expressed FLEX-GCaMP8s in cholinergic axons via virus injection into the horizontal diagonal band (HDB) of the BF in ChAT-cre mice as well as the red ACh sensor rACh1.0 in the VIS (Figure 1G). HDB was chosen due to its preferential projections to the VIS,[26–28] and we used rACh1.0 so that the sensor could be imaged simultaneously with GCaMP activity on red and green emission channels. We observed that BF axon activity was more transient and phasic than ACh activity, but both closely tracked locomotion periods and pupil size (Figure 1H). Thus, cholinergic axon activity is indeed tightly related to local ACh levels in the cortex, validating previous studies that have used axon activity alone to measure cholinergic tone[16,19] and ameliorating concerns about mechanisms that could potentially dissociate axonal action potentials from neuromodulator release.[29]

Changes in ACh levels surrounding periods of locomotion

As expected from previous studies,[20,21,30] ACh levels in the VIS increased during running. To more closely examine the timing, magnitude, and spatial heterogeneity of this increase, we measured and compared ACh activity during spontaneous running across three cortical areas: the VIS, somatosensory cortex (SS), and motor cortex (MO). When aligning ACh activity to all run onsets and offsets, we observed a similar general pattern across all three areas: a large increase right before run onset and a slower decline back to baseline after run offset (Figure 2A). To look at the timing of these onsets and offsets more closely, we restricted to isolated running periods with at least 15 s of quiet wakefulness before onsets or after offsets. Because recordings in the VIS, SS, and MO were not conducted simultaneously, we matched run periods by similar speed. We found that ACh activity began to increase 0.96 ± 0.1 s before run onset and declined after running with a decay constant of 5.8 ± 2.3 s (Figure 2B). While the onset timing was similar across all three brain areas, offset decay was significantly longer in the VIS than in the SS or MO. The timing of the initial ACh rise (within the first 5 s of running) was not dependent on run speed (Figure 2C). In contrast with a previous study[20] that found the same level of GRAB-ACh activation with both fast and slow run speeds, we observed a clear relationship between run speed and duration and the magnitude of ACh changes, with faster and longer run periods leading to larger ACh increases than slower and shorter run periods (Figure 2D). Run speed, duration, and magnitude of ACh response in the final 5 s of running did not have a significant effect on decay time (Figure 2E). Again, to compare ACh response magnitude across cortical areas, we matched run periods by run speed and duration and found significant differences across areas in both conditions (Figure 2F). Median peak GRAB-ACh ΔF/F was significantly higher in the SS and MO than in the VIS when matching by run speed and significantly higher in the MO than in the SS or VIS when matching by run duration. In summary, we observed the expected rapid increase of ACh preceding running and a slower decrease back to baseline right at run offset; the magnitude of this increase varied by run speed and duration and was relatively heterogeneous across cortical areas.

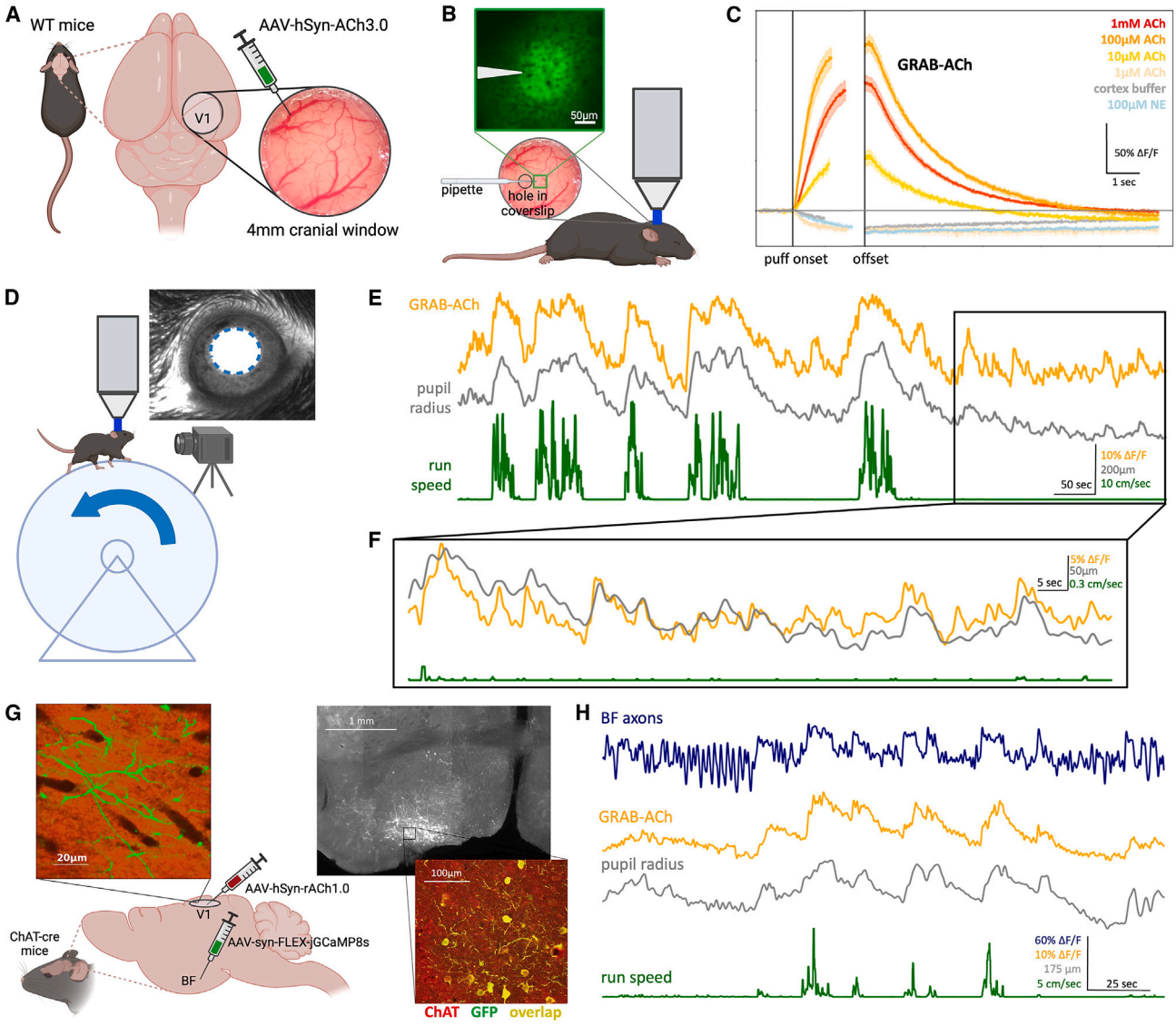


Figure 1. GRAB-ACh activity reveals spontaneous fluctuations in ACh levels that track behavior state and cholinergic axon activity

(A) Animal preparation. A cranial window was implanted over the VIS of wild-type (WT) mice, and GRAB-ACh3.0 was expressed via virus injection.

(B) For pharmacological validation experiments, a coverslip with a hole was implanted, through which a glass pipette could be inserted. Various concentrations of ACh were puffed onto the cortex while imaging was performed under isoflurane anesthesia.

(C) Puffs of varying concentrations of ACh resulted in increases in GRAB-ACh fluorescence, while injections of control substances (cortex buffer, norepinephrine) resulted in decreases in fluorescence. Puffs of 1 μM ACh also resulted in a decrease in fluorescence, consistent with an ambient ACh concentration above 1 μM. 3 mice total, *n* puffs in each group in the order listed = 10, 11, 5, 18, 15, 5; error bars represent standard error).

(D) Experimental setup for awake, behaving mouse experiments. The mouse was head fixed on a treadmill in a dark room underneath a two-photon microscope while treadmill speed and a video of the eye were recorded.

(E) Example traces of spontaneous ACh levels, pupil radius, and run speed while the mouse naturally transitioned between states. Run periods were accompanied by large increases in pupil size and ACh level (as indicated by GRAB-ACh fluorescence).

(F) A close up of the example trace in (E), highlighting the close correspondence of ACh with spontaneous pupil fluctuations during a period of stillness (quiescence).

(G) Animal preparation for axon imaging experiments. GCaMP8s was expressed in cholinergic axons by virus injection into the horizontal limb of the diagonal band (HDB) in the basal forebrain (BF) of ChAT-cre mice. A cranial window was implanted over the VIS, and rACh was expressed via virus injection. Verification of injection sites was subsequently done histologically.

(H) Example traces of spontaneous axon activity in comparison with simultaneous ACh, pupil radius, and run speed, in which correspondence between all four traces is apparent.

(A), (B), (D), and (G) were created with BioRender.

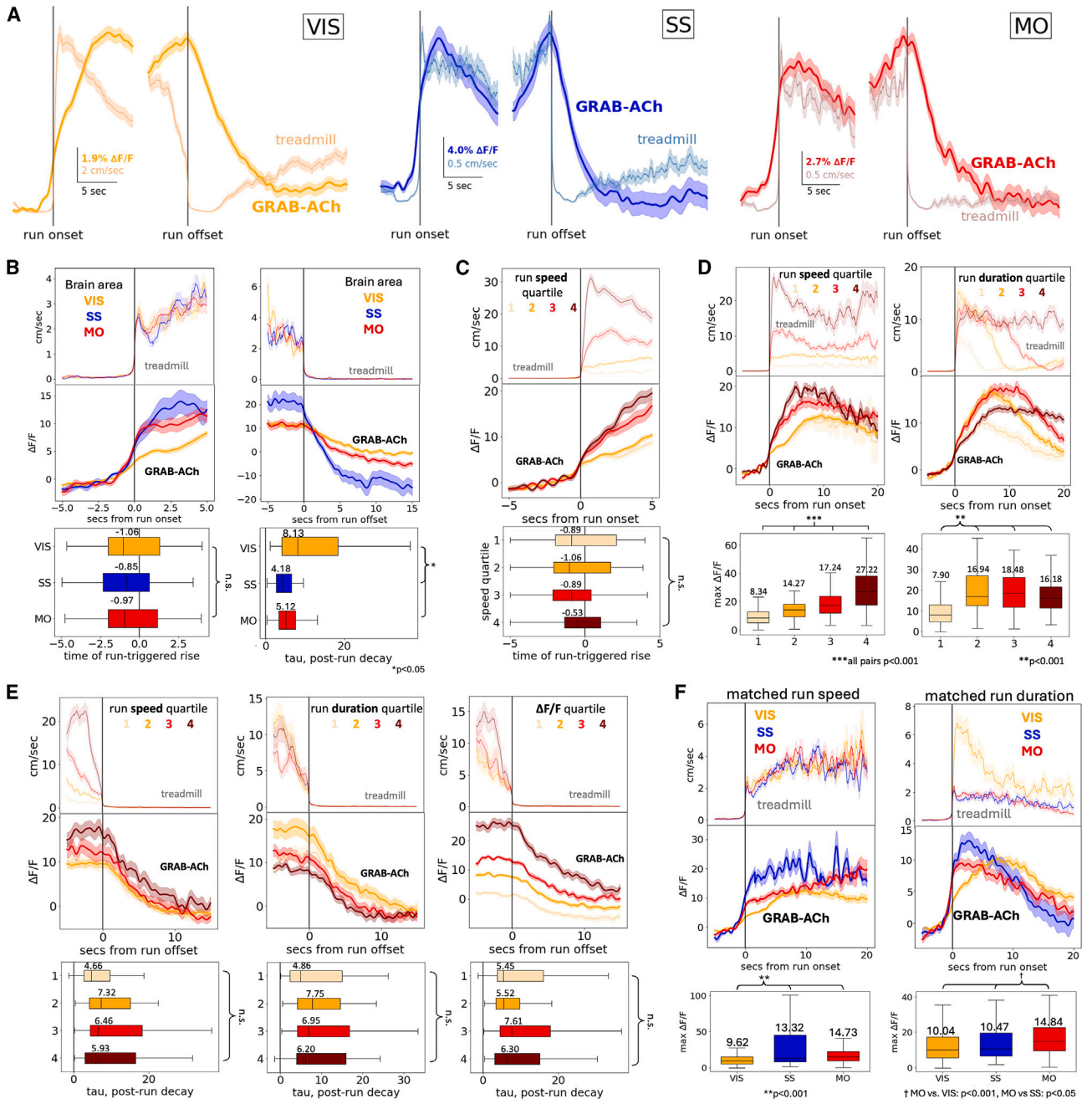


Figure 2. Locomotion is linked to large-scale, slowly decaying increases in ACh that vary by run magnitude and are somewhat heterogeneous across the dorsal cortex

(A) Mean run onset- and offset-triggered ACh traces for the VIS (n = 751 run periods from 53 regions of interest [ROIs] and 14 mice), SS (n = 289 runs from 26 ROIs and 6 mice), and MO (n = 206 runs from 19 ROIs and 5 mice).

(B) Mean run onset-triggered (n = 83 runs in each area) and offset-triggered (n = 58) ACh in the VIS, SS, and MO (matched by mean run speed in the first/last 5 s of running). Median differences across areas were significant for offset decay tau (overall p = 0.004, VIS vs. SS p = 0.001, VIS vs. MO p = 0.042, SS vs. MO p = 0.191, Kruskal-Wallis test) but not onset rise time (p = 0.718).

(C) In the VIS, mean run onset-triggered ACh binned by run speed in the first 5 s of running (n = 80–81 runs in each bin; see Figure S1 for a breakdown of run speeds). The median time of ACh rise in response to running was not significantly different across groups (p = 0.629).

(D) Mean run onset-triggered ACh in the VIS binned by mean speed in the entire run period (n = 100–101 runs in each in each bin) or total run duration (n = 99–100). In both conditions, the median peak ACh ΔF/F during running was significantly different across bins (overall p < 0.001; binning by run speed, all pairs p < 0.001; binning by run duration, 1 vs. 2, 1 vs. 3, and 1 vs. 4 p < 0.001; all other pairs not significant).

(legend continued on next page)

Changes in ACh levels surrounding dilation events during quiescence

As we have done previously,[1,19] we next excluded active periods to focus on the smaller fluctuations in pupil size that occurred outside of running. Certain previous studies have found that cholinergic axon activity[19] or ACh levels[26] are coherent with and precede changes in pupil size in the cortex, while others have found that pupil size explains relatively little variance in either GRAB-ACh or cholinergic axon activity.[16,30] We found that ACh fluctuations during quiet wakefulness were tightly tracked by small pupil dilations (Figure 3A). ACh activity was highly coherent with pupil size across brain regions, both during quiescence (Figure 3B) and when including periods of locomotion (Figure S3A). When we compared dilation periods of similar magnitude (Figure 3C), we found that larger dilation periods were accompanied by larger ACh increases as well as a shorter lag between ACh rise and dilation onset. Comparing dilation periods of similar magnitude across cortical areas (Figure 3D), our results suggest that dilation-linked ACh increases had a shorter lag in the VIS and that the magnitude of this increase differed significantly across all three areas. Thus, cortical ACh levels are correlated with and precede pupil dilation, although the timing and magnitude of this relationship may be heterogeneous across areas.

Temporal relationship between ACh and cholinergic axon activity

Previous studies have used calcium indicators expressed in axons as a proxy for changes in local ACh levels.[15,16,19] However, axon activity may be dissociated from extracellular neuromodulator levels, which depend on release probabilities and clearance kinetics. In order to directly measure the relationship between axon activity and fluctuations in ACh levels, we recorded cholinergic axon activity with GCaMP in the green channel and rACh in the red channel. As we have observed previously,[19] axon activity increased around running periods. Consistent with the hypothesis that changes in ACh levels are driven by local cholinergic axon activity, we observed that changes in axon activity preceded changes in ACh activity both at the onset and offset of running (Figure 4A). When we quantified this timing by examining run periods with no running in the 15 s before or after running, we found that axon activity began increasing 1.03 s before the rise in ACh, and GRAB-ACh fluorescence decayed 4.4 times as slow as the GCaMP axon activity after running, despite the similar off-kinetics of GCaMP and the rACh sensor (Figure 4B). Previous studies have shown similar responses of axons and neuromodulator sensors to related stimuli[31,32] and even in the same experimental paradigm.[33] However, the measurements reported here were performed simultaneously in the same animal, thereby enabling direct comparisons between cholinergic GCaMP axonal activity and local ACh release measured with fluorescent sensors. For example, these measurements demonstrate that ACh levels remain elevated for several seconds after axonal activity returns to baseline for large ACh transients, even after accounting for the decay timescale of the sensors themselves.

A related question is whether there is any effect of running magnitude on the accompanying axon or ACh activity. Previous evidence for this has been mixed, with one study suggesting that ACh levels do vary by run speed[21] and another study suggesting that neither ACh nor axon activity vary by run speed.[20] When binning by average run speed, we found that faster runs resulted in a larger increase in axon activity as well as ACh measured with the fluorescent sensor (Figure 4C). We found a similar pattern when binning by run duration: longer runs resulted in higher-magnitude axon and GRAB-ACh activity (Figure 4D). Interestingly, we noticed that the relative timing of these peaks in run-triggered axon and ACh activity was not consistent across runs of different durations. For longer run periods, the peak in ACh activity occurred significantly later than the peak in axon activity for that same run period (Figure 4E). Whereas axon activity typically peaked in about the first 4% of the run period, ACh activity did not reach its peak until just after the very end of the run period. Overall, these results suggest that ACh levels may continue to increase under steady-state axonal release, potentially due to saturation of clearance mechanisms.

Next, we examined simultaneously recorded axon and sensor dynamics with respect to pupil fluctuations during periods of quiescence. With respect to dilation onset, axon activity began increasing about 200 ms before ACh sensor activity (Figure 4F), and larger dilations resulted in a significantly larger magnitude of both axon and ACh sensor fluorescence (Figure 4G). This suggests that these smaller ACh changes associated with pupil size in the absence of running likely occur in a range where ACh is cleared rapidly, and, thus, tracks more closely with axonal activity. These smaller changes in ACh had a higher cross-correlation with pupil size than did axon activity (Figure S3D), indicating that spontaneous fluctuations in pupil size more closely track ACh changes than cholinergic axon activity. Furthermore, while rACh exhibited peak coherence with pupil size at lower frequencies (<0.1 Hz), both during quiescence and when accounting for running periods, cholinergic axon activity was most coherent with pupil size between 0.1 and 1Hz (Figures S3B and S3C). During quiescence we observed increased ~0.5 Hz coherence between axon and pupil, consistent with the timescale of pupil oscillations during resting periods, but lower frequencies encompassing the seconds-long running elevations in ACh were more prominent when running periods were included in the analysis (Figure S4).

(E) Mean run offset-triggered ACh in the VIS binned by mean run speed in the last 5 s of running, by total run duration, or by mean ΔF/F in the last 5 s of running (n = 50–51 runs in each bin). Median decay tau of ACh (fitted to the first 10 s after run offset) was not significantly different across bins in all three conditions (left to right, overall p = 0.181, 0.478, 0.475).

(F) Mean run onset-triggered ACh in each brain area matched by mean speed in the entire run period or total run duration (n = 150 runs in each area). Median peak ACh ΔF/F was significantly different across areas in both conditions (speed: overall $p < 0.001$, VIS vs. SS and VIS vs. MO $p < 0.001$, SS vs. MO $p = 0.389$; duration: overall $p < 0.001$, VIS vs. MO $p < 0.001$, SS vs. MO $p = 0.012$, VIS vs. SS $p = 0.110$).

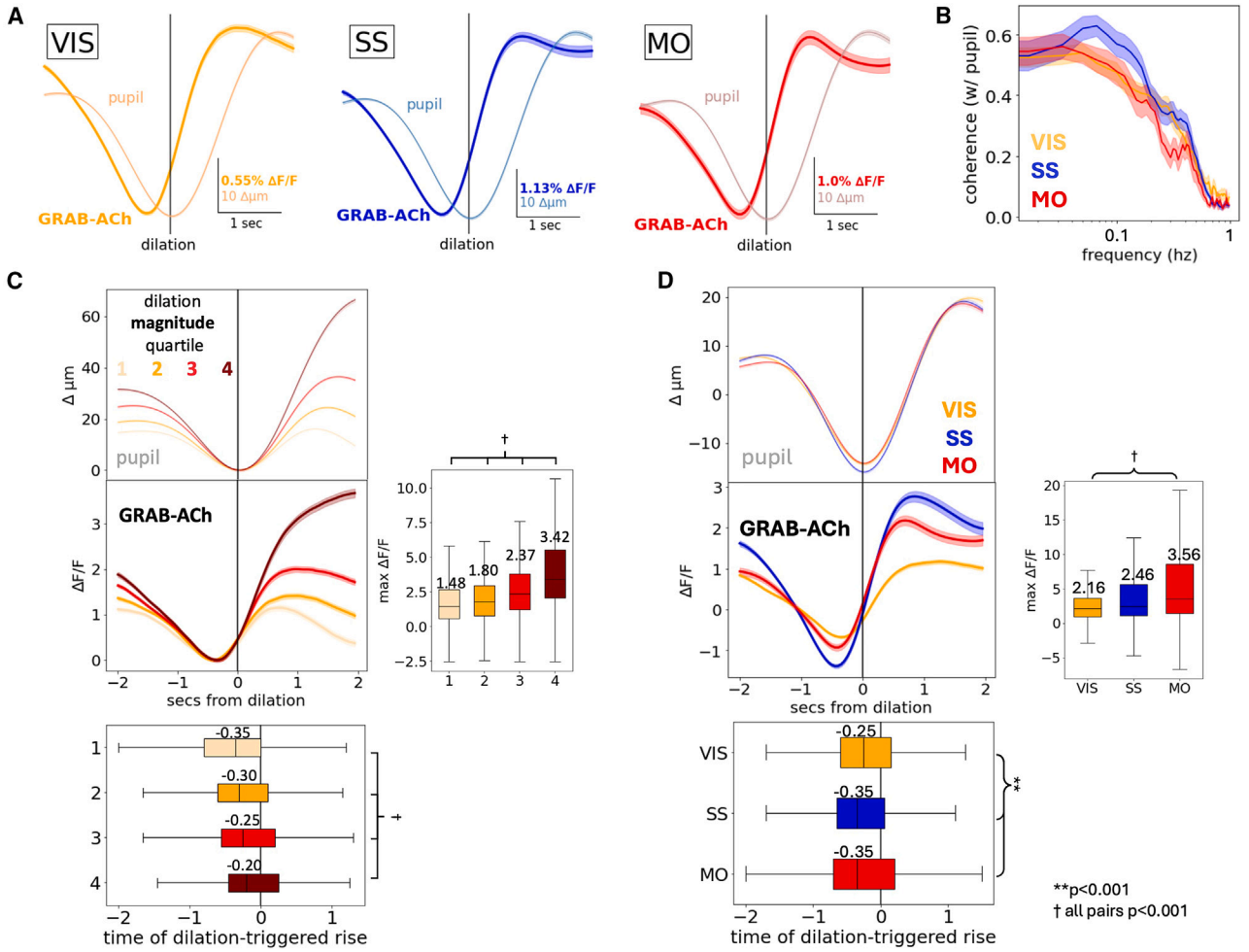


Figure 3. Pupil dilation during quiescence is linked to small-scale, rapid changes in ACh that vary by cortical area

(A) Dilation-triggered ACh traces during stillness for the VIS (n = 6,254 dilations from 44 ROIs; see Figure 1 for number of mice per brain area), SS (n = 4,335 from 20 ROIs), and MO (n = 3,069 from 15 ROIs). See Figure S2 for a characterization of pupil dynamics during quiescence.

(B) Mean coherence between pupil radius and ACh in different brain areas during quiescence. Across areas, coherences exhibited a broad peak at lower frequencies (<0.1 Hz).

(C) In the VIS, dilation-triggered ACh traces sorted by dilation magnitude (n = 1,563–1,564 dilations in each bin). Median time of dilation-triggered ACh rise was significantly different across bins (overall $p < 0.001$, all pairwise comparisons $p < 0.001$, Kruskal-Wallis test). Median peak ACh ΔF/F was also significantly different across bins (overall $p < 0.001$, all pairwise comparisons $p < 0.001$).

(D) Dilation-triggered ACh traces in the VIS, SS, and MO matched by dilation magnitude (n = 3,069 dilations in each area). Median time of dilation-triggered ACh rise was significantly different across areas (overall $p < 0.001$, VIS vs. SS and VIS vs. MO $p < 0.001$, SS vs. MO $p = 0.617$, Kruskal-Wallis test). Median peak ACh ΔF/F was also significantly different across areas (overall $p < 0.001$, all pairwise comparisons $p < 0.001$).

Spatial relationship between ACh and cholinergic axons and temporal modeling

We wondered whether our imaging configuration would provide sufficient resolution to observe spatial features of local ACh release surrounding cholinergic axons, which might be difficult, given that we only monitored the activity of axons in a 2D plane (i.e., we could potentially be missing axons above and below the field of view; Video S2). In each scan, we binned rACh-expressing pixels according to their distance from the nearest GCaMP-expressing axon in the field of view (six spatial bins; Figure 5A). To estimate axonal spiking, we deconvolved the axonal GCaMP fluorescence and applied a threshold to the resulting activity trace to restrict our analysis to larger isolated spikes (STAR Methods). Indeed, we found that stronger rACh responses were observed closer to the nearest axon (Figure 5B), with weaker responses from more distant pixels (monotonic decrease, Spearman's rank correlation coefficient = −0.066, $p < 0.001$). This result suggests that there is a small amount of spatial heterogeneity of ACh activity around cholinergic axons, consistent with a volume transmission mechanism (Video S2). It is important to note that, given the limitations of our imaging setup, we were only able to monitor the activity of axons in a 2D plane, and, thus, we could potentially be missing axons above and below the field of view. Moreover, it is possible that there could be other cholinergic axons close by but slightly outside the imaging FOV or even unlabeled (non-GCaMP-expressing) axons within the

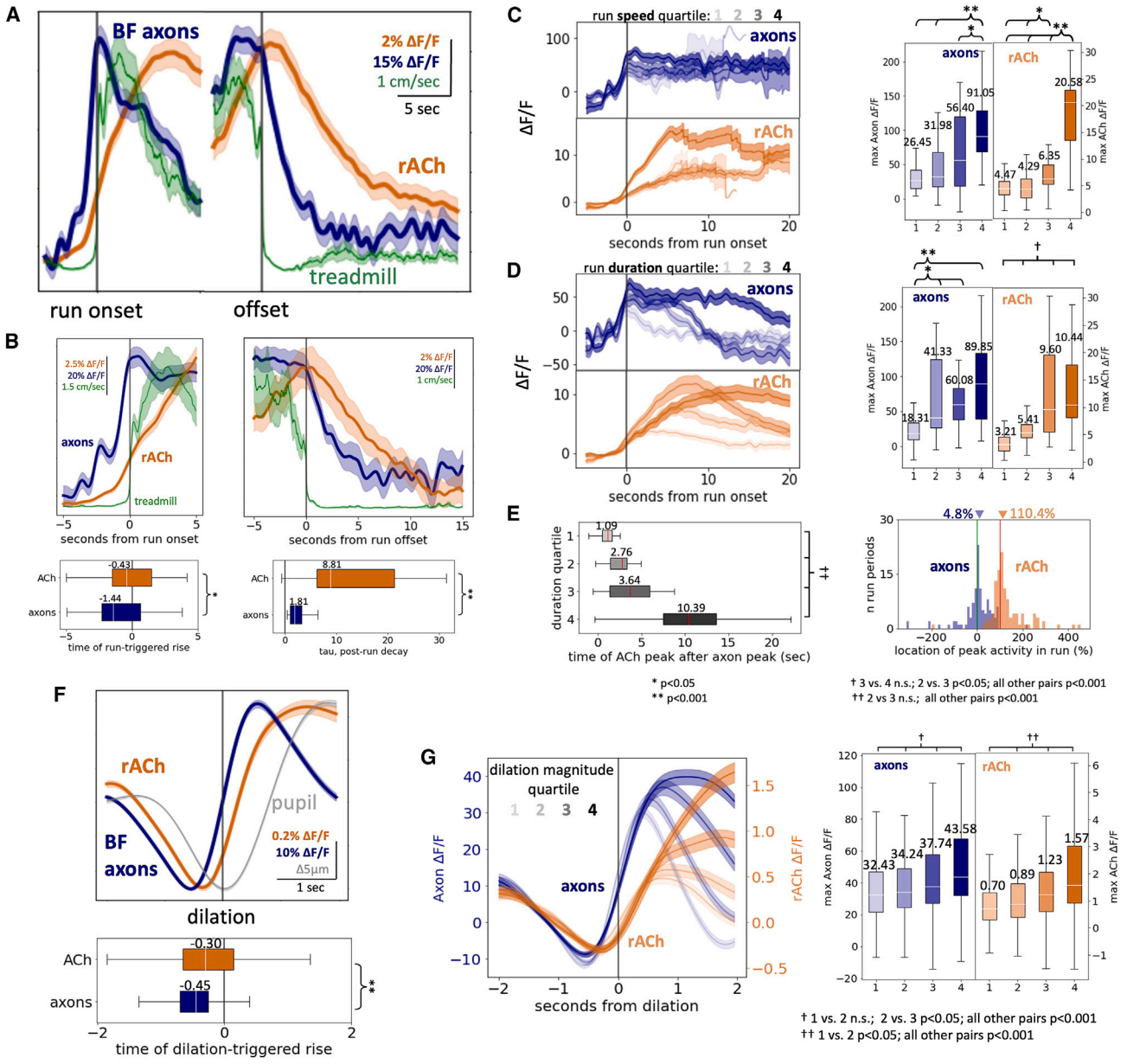


Figure 4. Cholinergic axon activity is accompanied by increases in local ACh levels that precede locomotion onset and pupil dilation

(A) Mean run onset- and offset-triggered ACh and axon traces (n = 194 run periods from 20 ROIs and 5 mice).

(B) Mean run onset-triggered (n = 78 runs from 19 ROIs) and offset-triggered (n = 38 runs from 19 ROIs) ACh and axon activity for run periods with no running 15 s before or after running were also excluded. Median time of increase in ACh activity was significantly later than time of increase in axon activity (p = 0.019, Kruskal-Wallis test). Median decay tau was also significantly faster for axon activity than ACh activity ($p < 0.001$).

(C) Mean run onset-triggered ACh and axon activity binned by mean run speed (n = 32–33 runs in each bin). There was a significant difference in median peak ΔF/F across bins for both axon and ACh activity (overall $p < 0.001$ for both groups; see figure for pairwise comparisons).

(D) Mean run onset-triggered ACh and axon activity binned by run duration (n = 32 runs in each bin). There was a significant difference in peak ΔF/F across bins for both axon and ACh activity (overall axons $p < 0.001$ for both groups; see figure for individual pairwise comparisons).

(E) Left: time of peak ACh ΔF/F after peak axon ΔF/F for the run bins in (D). There was a significant difference across bins (overall $p < 0.001$; see figure for pairwise comparisons). Right: histogram of location of peak axon vs. peak ACh peri-run ΔF/F, expressed as a percentage of run (where 0% is peak occurring at run onset and 100% is peak occurring at run offset). Axon traces peaked early, at 4.8% of the run (median across axons), whereas ACh signals peaked later, at 110.4% of the run.

(F) Dilation-triggered ACh and axon traces during stillness (n = 1,525 dilation periods from 16 ROIs). Median time of increase in ACh activity was significantly later than time of increase in axon activity ($p < 0.001$).

(G) Dilation-triggered ACh and axon traces binned by dilation magnitude (n = 381–382 dilations in each bin). There was a significant difference in median peak ΔF/F for both axon activity (overall $p < 0.001$; see figure for pairwise comparisons) and ACh activity (overall $p < 0.001$; see figure for pairwise comparisons).

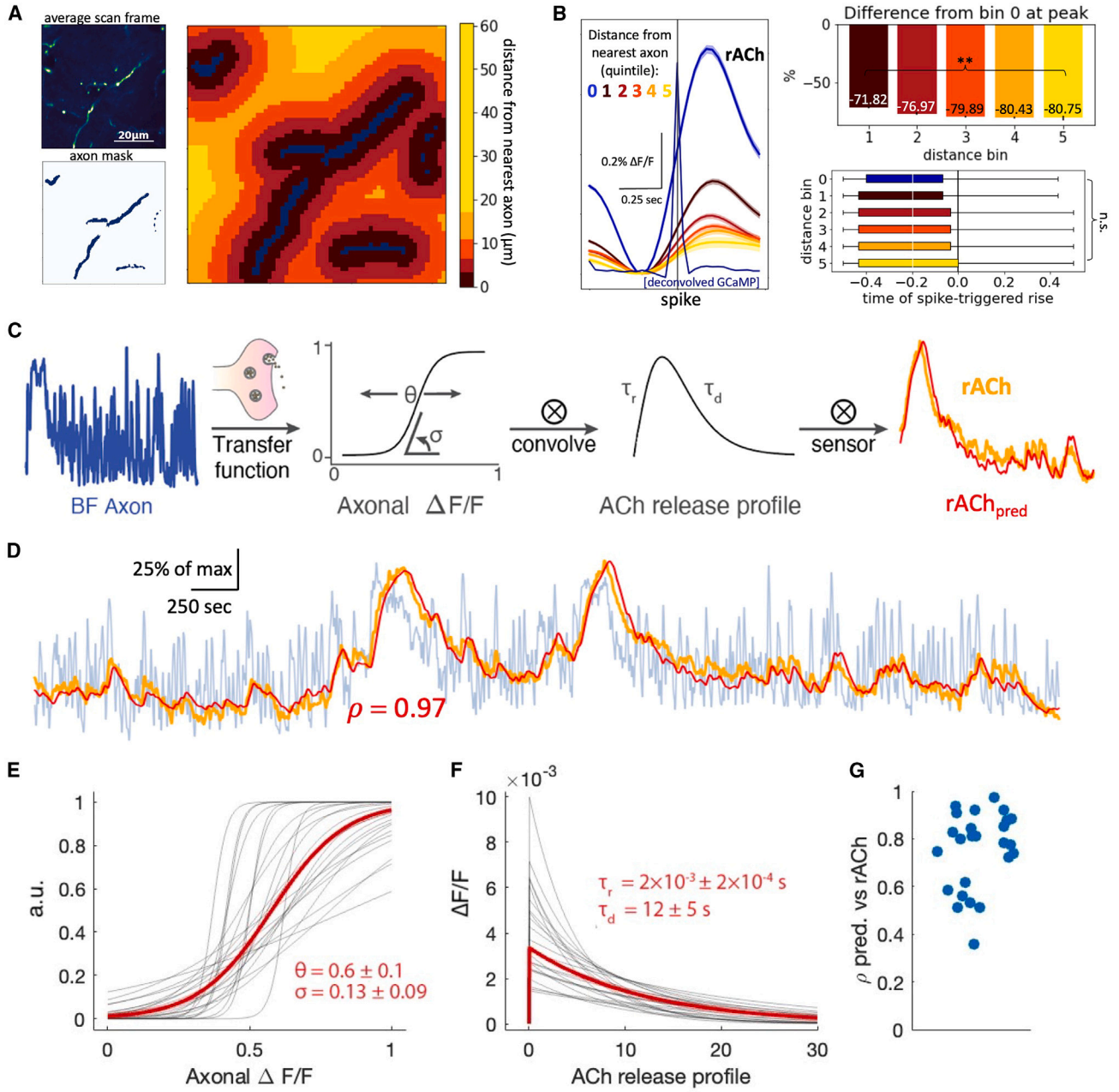


Figure 5. ACh activity decreases at distances farther from an axon and can be predicted temporally from cholinergic axon activity

(A) For one example scan, average scan frame, ROI mask representing location of axons, and a map representing pixel-wise distance from nearest axon as defined by the axon mask.

(B) Left: mean spike-triggered ACh ΔF/F for all pixels in all recordings (n = 6,747 spikes from 23 recordings, inter-spike interval = ±0.5 s) binned by distance from nearest axon, where bin 0 represents pixels containing an axon mask, and one "spike" represents thresholded activity of deconvolved GCaMP. Top right: for each bin, mean ΔF/F value at peak activity in bin 0 expressed as percent change from bin 0. The rank order was significant (Spearman rank order correlation coefficient = −0.066, $p < 0.001$). Bottom right: time of spike-triggered rACh rise. Median value for every bin is −0.2 s.

(C) Prediction model workflow going from recorded cholinergic axon activity to predicted rACh signal with Θ, σ, τ_r, and τ_d as the fitted parameters.

(D) An example segment of recorded axon activity (blue), recorded rACh activity (orange), and predicted rACh activity (red). Shown is Pearson correlation ρ = 0.97 between recorded and predicted rACh.

(E) Sigmoidal activation function from axon activity to ACh release for each recording (black, n = 25 recordings from 5 mice) and averaged over all recordings (red).

(F) ACh release functions for each recording (black) and averaged over all recordings (red).

(G) Pearson correlation between empirical and predicted rACh for each recording.

FOV that could be contributing to the measured ACh release. Given the fact that we still observed a clear spatial relationship between ACh levels and axon proximity despite these potential confounds, our result may actually represent a lower bound on the amount of spatial structure in ACh release.

In order to quantify the extent to which GRAB-ACh fluorescence fluctuations could be predicted from axon activity, we built a simple model with a sigmoidal function relating GCaMP axonal activity to ACh release and an exponential rise (τ_r) and decay[34] (τ_d; Figure 5C) for ACh. We found that this simple dynamic model was able to use axonal fluorescence (Figure 5D, blue) to generate a prediction for rACh (Figure 5D, red) that closely resembled empirically measured rACh (Figure 5D, orange; Pearson correlation $\rho = 0.97$ between observed and predicted rACh). Overall, we found that the optimization algorithm predicted a slightly right-shifted sigmoidal (i.e., more calcium required for a response; Figure 5E) and a rapid release ($\tau_r \sim 0$) with decay over seconds ($\tau_r \sim 12$s; Figure 5F). Across all recordings, the algorithm could accurately predict rACh fluorescence ($\bar{\rho} = 0.8 \pm 0.2$; Figure 5G) and explain a significant amount of the rACh variance ($\bar{\rho}^2 = 0.6 \pm 0.2$). This algorithm provides a simple method to use recordings of axonal fluorescence to predict what the GRAB-ACh sensor would report. More generally, it provides a strong validation for existing and future studies using cholinergic axon activity as a proxy for measuring local ACh levels in the cortex, with the caveat that ACh levels and axon activity levels may not be linearly related for large cholinergic events.

Deconvolution of GRAB-ACh sensors

Having confirmed that cortical ACh levels track state-related variables and cholinergic axon activity, we wanted to better understand the actual time course of ACh fluctuations independent of GRAB-ACh sensor kinetics. We estimated the time course of [ACh] by deconvolving the GRAB-ACh impulse response measured *in vitro* from the fluorescence trace (Figure 6A; more details in STAR Methods). Applying this deconvolution approach to ACh puffs (as in Figure 1) yielded the expected faster rise and return to baseline of [ACh] compared with the GRAB-ACh sensor (Figure 6B). Large-scale ACh changes such as those surrounding running were too slow to be substantially affected by the sensor off time (Figures 6C and 6D). However, deconvolution emphasized that the rapid clearance of smaller changes in ACh levels (such as those associated with pupil dilation in the absence of running) may be obscured by the off kinetics of the GRAB-ACh sensor (Figures 6E–6H). By chance, the sensor kinetics for GCaMP and rACh were similar enough that the lag between them was similar for deconvolved and non-deconvolved traces (Figure 6G). It should be noted that [ACh] appears to rise before the spike onset rather than after; we expect that this may be due to spiking activity excluded by our moderate thresholding of the GCaMP activity trace and is also accentuated by low-pass filtering of the GRAB-ACh trace.

Prediction of GRAB-ACh activity from locomotion and pupil size

Finally, given our results indicating a close relationship between ACh and state-related behavior variables, we wondered how accurately we could predict ACh levels from these variables. To this end, we trained a long short-term memory (LSTM) recurrent neural network to predict GRAB-ACh fluctuations from the recorded behavioral dynamics: pupil size and running speed and their temporal derivative (Figure 7A). The statistical model was able to reproduce the GRAB-ACh from training recordings ($\bar{\rho} = 0.73 \pm 0.01$; Figure 7B), explaining over 50% of the GRAB-ACh variance ($\bar{\rho}^2 = 0.57 \pm 0.01$), both during high-arousal states marked by locomotion as well as quiescent states (Figure 7A right). As expected, the model's explanatory capacity improved with an increase in correlation between the behavioral variables and GRAB-ACh signal, with model accuracy routinely exceeding behavioral signal by itself (Figure 7C). Importantly, the model was able to correctly predict GRAB-ACh from behavior in unseen data (held-out validation dataset) at a performance matching training data (Figure 7D, second from left; explained variance $\bar{\rho}^2 = 0.57 \pm 0.23$). While there is likely to be some variability from setup to setup, in general, pupil size and running speed are fairly straightforward variables to measure. The good predictive performance of the model on the held-out data suggests that the model could potentially generalize to predict ACh levels in other head-fixed mouse experiments with minimal retraining. As both behavioral variables are used as proxies for arousal, we wanted to contrast their explanatory power. Using either behavioral metric (pupil or locomotion speed) independently led to a decrease in signal accuracy compared to using both variables, as did using a circularly shuffled dataset or only periods outside of running (Figure 7D; pupil $p = 1 \times 10^{-5}$, speed $p = 1 \times 10^{-7}$, null $p = 2 \times 10^{-34}$, still $p = 3 \times 10^{-8}$, paired t test). When looking at data across brain regions, we found that the variability in model performance within a brain area was larger than the variability between areas (Figure 7E), and models trained and tested within or across areas performed equally well, again consistent with the general applicability of the model. It should be noted that this model predicts GRAB-ACh sensor activity rather than absolute levels of [ACh], but because the deconvolution from GRAB-ACh activity to [ACh] is a simple linear transformation, the deconvolution could easily be applied to the predicted GRAB-ACh fluorescence as a downstream processing step. Overall, we demonstrated that behavioral variables used to demarcate changes in brain state can also be used to predict ACh levels in the cortex.

DISCUSSION

Our study leverages recent advances in genetically encoded fluorescent indicators to investigate the question of when and where ACh is available during spontaneous, fast brain-state transitions. Consistent with previous work,[20,21,30] we found that transitions from quiet wakefulness to the high-arousal state during locomotion resulted in a large increase in ACh that outlasted running periods by several seconds. Also consistent with previous results,[26] small ACh increases occurred around a half second before the onset of spontaneous epochs of pupil dilation that occur when animals are sitting quietly.

Previous studies have presented mixed results regarding whether there is a linear relationship between run speed and ACh response[20,21]; that is, whether faster run speeds are accompanied by larger increases in ACh activity than slower run

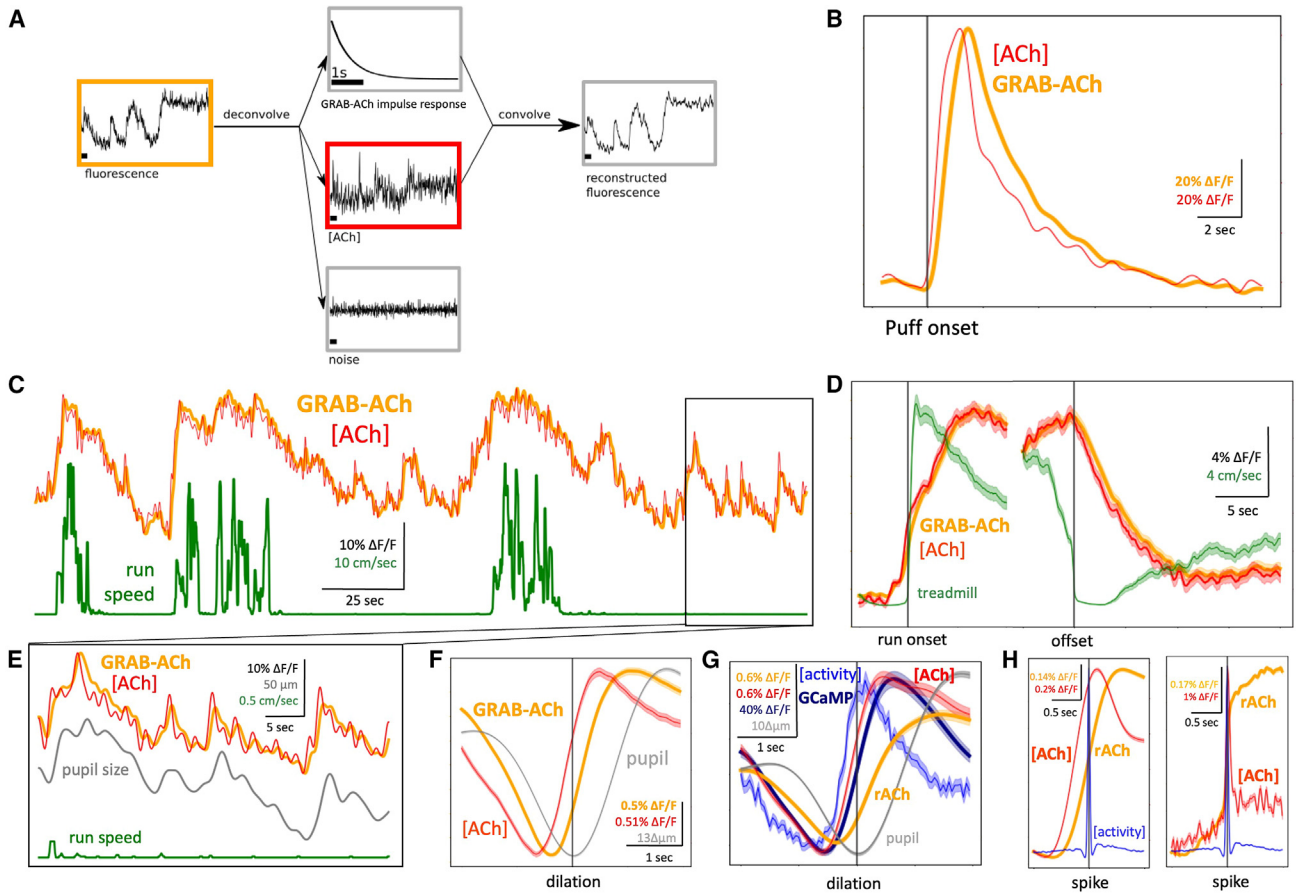


Figure 6. For small and fast events, GRAB-ACh deconvolution can help reveal rapid ACh kinetics

(A) Schematic illustrating the causal deconvolution filter.

(B) A single 100 μM ACh injection-triggered GRAB-ACh trace with the deconvolved ACh trace overlaid. As expected, deconvolved ACh reaches a peak slightly before GRAB-ACh ΔF/F.

(C) An example trace of GRAB-ACh in the VIS, along with deconvolved ACh and running trace.
Deconvolution does not substantially alter the large-scale changes of ACh seen around running.

(D) Run onset- and offset-triggered GRAB-ACh and deconvolved ACh for all running periods in VIS recordings. Deconvolution does not substantially alter the timing of ACh increases before running or decay after running.

(E) Close up of a quiescence period in the example traces in (C), with the addition of pupil size. Deconvolution reveals sharper, more phasic activity of ACh correlated with spontaneous pupil size fluctuations.

(F) Dilation-triggered GRAB-ACh and deconvolved ACh traces for all dilation periods in VIS recordings. Deconvolved ACh rises and peaks sooner than is indicated by GRAB-ACh fluorescence. Differences in ΔF/F magnitude between deconvolved and non-deconvolved traces should not be interpreted.

(G) Dilation-triggered rACh and axon activity (GCaMP) with deconvolved ACh and axon activity traces. After deconvolution, the order of axon activity and ACh activity revealed in previous experiments is preserved.

(H) Spike-triggered GRAB-ACh and deconvolved ACh for all simultaneous axon-GRAB-ACh recordings (n = 45,669 spikes; left, 1 Hz low pass filtered; right, unfiltered). One "spike" here represents thresholded activity of deconvolved GCaMP fluorescence. Again, deconvolved ACh rises and peaks sooner than what is indicated by GRAB-ACh fluorescence.

speeds. Here, we found that faster and longer run periods do indeed result in a larger ACh response. Differences in findings between our study and previous studies could be due to the exact time point where ACh is measured relative to running speed, since we find that the peak in ACh response occurs late in the running period. Similarly, we found a clear positive association between pupil dilation magnitude and ACh response in our data, a point on which some studies differ.[26,30]

We sought to examine the heterogeneity of ACh release and clearance across several areas of the dorsal cortex. While the BF is often referred to as a monolithic structure, evidence suggests that it is actually somewhat modular with segregated projections to the cortex.[35,36] The VIS receives a large proportion of its cholinergic projections from HDB, while the auditory cortex is preferentially targeted by the posterior nucleus basalis (NB) and the SS by the anterior NB.[27] We found that the magnitude and timing of both run- and dilation-evoked ACh differed significantly across brain areas. Interestingly, we found higher ACh release at run onset in the MO than in the VIS, which is congruent with a recent study[30] that also found higher ACh activity in anterior

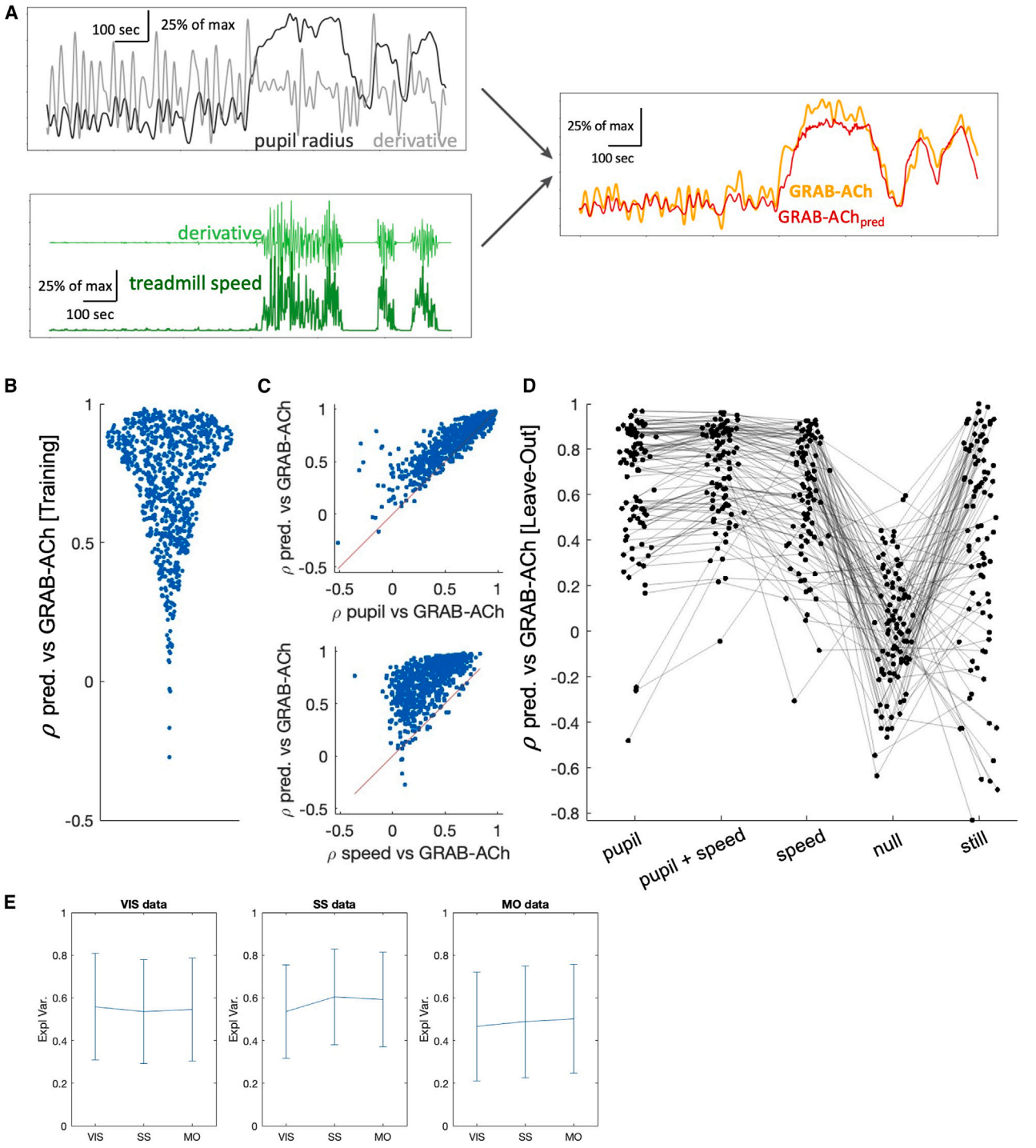


Figure 7. ACh levels can be predicted from pupil size and locomotion speed

(A) A model was constructed from pupil and treadmill measurements and their derivatives. Shown are example traces of pupil radius/derivative and treadmill speed/derivative and the accompanying observed GRAB-ACh activity with predicted GRAB-ACh.

(B) Correlation between observed and predicted GRAB-ACh for each training sample (n = 834 2-min-long epochs over 104 recordings and 24 mice).

(C) Top: for each training sample, correlation between pupil size and observed GRAB-ACh versus correlation between predicted and observed GRAB-ACh (81.4% above the diagonal). Bottom; the same comparison but instead with treadmill speed (97.8% above the diagonal).

(legend continued on next page)

than posterior cortical areas at locomotion onset. In addition, we found that small-scale ACh increases—linked to pupil dilation—were larger in VIS than in the more anterior cortical areas and occurred with less of a delay relative to dilation.

To assess whether there was a direct relationship between local ACh levels and cholinergic axon activity, we performed simultaneous measurements of cholinergic axons using GCaMP and the accompanying changes in ACh levels using a fluorescent ACh sensor along with behavior. While other studies have suggested that axon activity and neuromodulator sensors convey similar information,[31–33] we directly confirmed this by measuring both axons and ACh *in vivo* simultaneously in the same FOV. Because many studies have used axon activity as a proxy for ACh levels,[15–17,19] these results are relevant for bridging across previous literature. Broadly, we confirmed that local ACh levels do indeed correspond to axon activity, as expected, and that the longer-lasting increases in ACh seen during running are much slower to decay than the rapid, phasic increases linked to dilation. This suggests that, in line with pharmacological studies indicating the rapid action of acetylcholinesterase,[37] small amounts of ACh are cleared quickly, but larger increases may transiently saturate clearance mechanisms, leading to longer decays.[38] Consistent with this scenario, we found that the timing and magnitude of small increases of ACh around pupil dilation vary depending on dilation magnitude.

Spatially, we found that ACh availability decreases with increasing distance from axons, consistent with volume transmission of ACh. Although it was previously debated whether ACh can escape the synapse and influence receptors on neighboring processes not directly synapsing onto a cholinergic axon,[38,39] our results are in line with the growing consensus that ACh does indeed escape the synapse. While ACh spillover has been detected using a variety of methods, including GRAB-ACh in an *in vitro* preparation,[40] our finding that the magnitude of cholinergic spike-triggered ACh activity displays a distance-dependent relationship is direct confirmation of volume transmission on a fine temporal and spatial scale *in vivo*. Overall, our simultaneous rACh/axon recordings support the notion that cholinergic axon activity can indeed be used as an accurate proxy for changes in local ACh levels, which is reassuring given the many existing studies that have used ChAT-expressing axons as a proxy for ACh.

Given the close correspondence of ACh activity with axon activity and pupil size or locomotion, we wanted to quantify the extent to which it is possible to predict ACh levels from these variables. Predictive modeling revealed that cholinergic axon activity and behavioral state variables, especially pupil size, can be used for accurate prediction of cortical ACh levels, explaining the majority of variance on held-out data. Pupil and treadmill activity is routinely recorded in many head-fixed mouse experiments, and so this model may be useful for inferring ACh fluctuations in the absence of ground-truth training data. Throughout this study, we uncovered a close link between pupil size and cortical ACh levels. We do not interpret this as cortical ACh controlling pupil size but, rather, that both cortical ACh levels and the pupil have a common upstream input. The anatomical pathway by which this could occur is not clear, but it is possible that fluctuations in cortical ACh delivered via the BF may be occurring downstream of projections from the locus coeruleus (LC) to the BF,[41,42] as there are several putative pathways by which the LC could control pupil size.[43] Although evidence showing that LC axonal activity precedes BF activity before pupil dilation[19] supports this possibility, further studies are needed.

In addition to the cholinergic system,[15,19,26] pupil size has also been shown to be correlated with the activity of other neuromodulatory systems, including the adrenergic[19,44] and serotonergic[45] systems. It is not uncommon for pupil size to be used as an index for activation of the adrenergic LC system in humans.[46–48] While our results should not be interpreted to mean that ACh is the only or best neuromodulator correlated with pupil size, they do emphasize that caution should be taken when using pupil size as a proxy exclusively for noradrenergic activity. Furthermore, we believe our study underscores the potential usefulness of pupillometry for inferring changes in ACh, as has been proposed elsewhere.[49]

Our study does not differentiate between extracellular ACh originating from BF projections and from cortical cholinergic interneurons. The primary source of cortical ACh has long been accepted to be the BF,[50] but more recently it has been discovered that VIP/ChAT-positive interneurons also release ACh in the cortex.[51,52] While some cholinergic transmission from these cells has been reported in layer I,[51] where most of our imaging occurred, the cell bodies of these sparsely expressed bipolar neurons primarily reside in layer II/III[53] and typically have downward-projecting axons.[54] While we cannot rule out that ACh from these cells is not influencing the ACh dynamics we observe here, given the projection pattern and sparse expression of ChAT interneurons, as well as the close correspondence we observed between BF axons and local ACh signal, it is reasonable to assume that the ACh signal we recorded was largely dominated by BF-derived ACh.

Our study provides a bridge between previous work using cholinergic axonal GCaMP imaging and recently developed genetically encoded fluorescent sensors and provides additional validation of GRAB-ACh for investigating the question of when and where ACh is available in the brain. The deconvolution approach we demonstrate here may be useful for other situations where it is desirable to estimate the time course of changes in levels of a ligand while removing the temporal smoothing due to slow sensor kinetics. Overall, these experiments contribute to a growing understanding of the spatiotemporal scale of ACh availability in the cortex in relation to state-related behavioral variables. This understanding is essential for future work focused on studying the downstream effects of ACh in changes on cortical excitability and information processing during perception and behavior.

(D) For the leave-out data (n = 93 samples), correlations between observed and predicted GRAB-ACh when using as prediction variables (left to right): pupil size alone (paired t test compared to pupil + speed, $p = 1 \times 10^{-5}$), pupil size and treadmill speed together, treadmill speed alone ($p = 1 \times 10^{-7}$), pupil size and treadmill speed with circularly shuffled data ($p = 2 \times 10^{-34}$), and pupil size and treadmill speed together during only periods of quiescence (outside of running) ($p = 3 \times 10^{-8}$).
(E) Select data from the VIS, SS, or MO were used to train an LSTM model in order to predict GRAB-ACh activity in the other brain regions.

Limitations of the study

Although the current study discovered small differences in the timing and magnitude of brain state-dependent ACh level changes across brain areas, these recordings were not conducted simultaneously, and therefore the noted differences could potentially be partly due to animal-to-animal variability. Our analysis that found that ACh availability decreases with increasing distance from axons is also limited by having imaged only axons in a single plane; this finding could be verified by conducting imaging in a 3D volume. Finally, we were unable to conclusively demonstrate that the ACh measured in this study originated from the BF and not from cortical interneurons. The source of ACh could be differentiated by measuring cortical GRAB-ACh activity after silencing ACh-expressing ChAT interneurons.

RESOURCE AVAILABILITY

Lead contact

Requests for further information, resources, and reagents should be directed to and will be fulfilled by the lead contact, Jacob Reimer (reimer@bcm.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Fluorescence data have been deposited at the DANDI archive and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We would like to thank the Andreas Tolias lab for providing resources and equipment. E.N. was supported by NIH NRSA Fellowship F31NS122428. C.S. was supported by NIH NRSA Fellowship F31NS132465. Y.L. was supported by The National Science Fund for Distinguished Young Scholars and NIH BRAIN Initiative (1U01NS103558, 1U01NS113358, and 1U01NS12 0824). M.J.M. was supported by NIDCD awards R01 DC017797 and R03 DC015618. J.M.S. was supported by the National Health and Medical Research Council (GNT1193857) and the Bellberry Foundation. J.R. was supported by NIH awards 1R34NS132045, 1RF1NS128901, and R56 AG080735.

AUTHOR CONTRIBUTIONS

Conceptualization, J.R.; methodology, B.R.M., R.G.L., and J.R.; software, C.S.; formal analysis, E.N.; investigation, E.N., N.Z., Z.H.M., F.A.B., and G.L.; resources, Y.L., M.H., M.J.M., J.M.S., and J.R.; data curation, E.N. and M.H.; writing – original draft, E.N.; writing – review & editing, B.R.M., R.G.L., M.J.M., J.M.S., and J.R.; visualization, E.N.; supervision, M.J.M., J.M.S., and J.R.; funding acquisition, E.N., C.S., Y.L., M.J.M., J.M.S., and J.R.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114808.

Received: January 29, 2024
Revised: July 18, 2024
Accepted: September 13, 2024

REFERENCES

1. Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., and Tolias, A.S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. Neuron *84*, 355–362.
2. McGinley, M.J., David, S.V., and McCormick, D.A. (2015). Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron *87*, 179–192.
3. Polack, P.-O., Friedman, J., and Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nat. Neurosci. *16*, 1331–1339.
4. Bennett, C., Arroyo, S., and Hestrin, S. (2013). Subthreshold Mechanisms Underlying State-Dependent Modulation of Visual Responses. Neuron *80*, 350–357.
5. Aydın, Ç., Couto, J., Giugliano, M., Farrow, K., and Bonin, V. (2018). Locomotion modulates specific functional cell types in the mouse visual thalamus. Nat. Commun. *9*, 4882.
6. Mineault, P.J., Tring, E., Trachtenberg, J.T., and Ringach, D.L. (2016). Enhanced Spatial Resolution During Locomotion and Heightened Attention in Mouse Primary Visual Cortex. J. Neurosci. *36*, 6382–6392.
7. Erisken, S., Vaiceliunaite, A., Jurjut, O., Fiorini, M., Katzner, S., and Busse, L. (2014). Effects of Locomotion Extend throughout the Mouse Early Visual System. Curr. Biol. *24*, 2899–2907.
8. Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. Neuron *86*, 740–754.
9. Dadarlat, M.C., and Stryker, M.P. (2017). Locomotion Enhances Neural Encoding of Visual Stimuli in Mouse V1. J. Neurosci. *37*, 3764–3775.
10. Tantirigama, M.L.S., Zolnik, T., Judkewitz, B., Larkum, M.E., and Sachdev, R.N.S. (2020). Perspective on the Multiple Pathways to Changing Brain States. Front. Syst. Neurosci. *14*, 23.
11. Metherate, R., Cox, C.L., and Ashe, J.H. (1992). Cellular bases of neocortical activation: modulation of neural oscillations by the nucleus basalis and endogenous acetylcholine. J. Neurosci. *12*, 4701–4711.
12. Williams, S.R., and Fletcher, L.N. (2019). A Dendritic Substrate for the Cholinergic Control of Neocortical Output Neurons. Neuron *101*, 486–499.e4.
13. Gill, D.F., and Hansel, C. (2020). Muscarinic modulation of SK2-type K^+ channels promotes intrinsic plasticity in L2/3 pyramidal neurons of the

mouse primary somatosensory cortex. eNeuro *7*, 2020. https://doi.org/10.1523/ENEURO.0453-19.2020.

14. Xiang, Z., Huguenard, J.R., and Prince, D.A. (1998). Cholinergic Switching Within Neocortical Inhibitory Networks. Sci. Am. Assoc. Adv. Sci. *281*, 985–988.
15. Nelson, A., and Mooney, R. (2016). The Basal Forebrain and Motor Cortex Provide Convergent yet Distinct Movement-Related Inputs to the Auditory Cortex. Neuron *90*, 635–648.
16. Collins, L., Francis, J., Emanuel, B., and McCormick, D.A. (2023). Cholinergic and noradrenergic axonal activity contains a behavioral-state signal that is coordinated across the dorsal cortex. Elife *12*, e81826.
17. Eggermann, E., Kremer, Y., Crochet, S., and Petersen, C.C.H. (2014). Cholinergic Signals in Mouse Barrel Cortex during Active Whisker Sensing. Cell Rep. *9*, 1654–1660.
18. Zou, J., Willem De Gee, J., Mridha, Z., Trinh, S., Erskine, A., Jing, M., Yao, J., Walker, S., Li, Y., McGinley, M., and Hires, S.A. (2024). Goal-directed motor actions drive acetylcholine dynamics in sensory cortex. Elife. https://doi.org/10.7554/eLife.96931.1.
19. Reimer, J., McGinley, M.J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D.A., and Tolias, A.S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat. Commun. *7*, 13289.
20. Yogesh, B., and Keller, G.B. (2024). Cholinergic input to mouse visual cortex signals a movement state and acutely enhances layer 5 responsiveness. Elife *12*, RP89986.
21. 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.
22. Vinson, P.N., and Justice, J.B. (1997). Effect of neostigmine on concentration and extraction fraction of acetylcholine using quantitative microdialysis. J. Neurosci. Methods *73*, 61–67.
23. Parikh, V., Kozak, R., Martinez, V., and Sarter, M. (2007). Prefrontal Acetylcholine Release Controls Cue Detection on Multiple Timescales. Neuron *56*, 141–154.
24. Yogesh, B., and Keller, G.B. (2023). Cholinergic Input to Mouse Visual Cortex Signals a Movement State and Acutely Enhances Layer 5 Responsiveness. Preprint at bioarXiv. https://doi.org/10.1101/2023.06.07.543871.
25. Bulumulla, C., Krasley, A.T., Cristofori-Armstrong, B., Valinsky, W.C., Walpita, D., Ackerman, D., Clapham, D.E., and Beyene, A.G. (2022). Visualizing synaptic dopamine efflux with a 2D composite nanofilm. Elife *11*, e78773.
26. Robert, B., Kimchi, E.Y., Watanabe, Y., Chakoma, T., Jing, M., Li, Y., and Polley, D.B. (2021). A functional topography within the cholinergic basal forebrain for encoding sensory cues and behavioral reinforcement outcomes. Elife *10*, e69514.
27. Kim, J.-H., Jung, A.-H., Jeong, D., Choi, I., Kim, K., Shin, S., Kim, S.J., and Lee, S.-H. (2016). Selectivity of Neuromodulatory Projections from the Basal Forebrain and Locus Ceruleus to Primary Sensory Cortices. J. Neurosci. *36*, 5314–5327.
28. Huppé-Gourgues, F., Jegouic, K., and Vaucher, E. (2018). Topographic Organization of Cholinergic Innervation From the Basal Forebrain to the Visual Cortex in the Rat. Front. Neural Circuits *12*, 19.
29. Wu, L.-G., and Saggau, P. (1997). Presynaptic inhibition of elicited neurotransmitter release. Trends Neurosci. *20*, 204–212.
30. Lohani, S., Moberly, A.H., Benisty, H., Landa, B., Jing, M., Li, Y., Higley, M.J., and Cardin, J.A. (2022). Spatiotemporally heterogeneous coordination of cholinergic and neocortical activity. Nat. Neurosci. *25*, 1706–1713.
31. Lovett-Barron, M., Kaifosh, P., Kheirbek, M.A., Danielson, N., Zaremba, J.D., Reardon, T.R., Turi, G.F., Hen, R., Zemelman, B.V., and Losonczy, A. (2014). Dendritic Inhibition in the Hippocampus Supports Fear Learning. Science *343*, 857–863.
32. Allard, S., and Hussain Shuler, M.G. (2023). Cholinergic Reinforcement Signaling Is Impaired by Amyloidosis Prior to Its Synaptic Loss. J. Neurosci. *43*, 6988–7005.
33. Kim, H.R., Malik, A.N., Mikhael, J.G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-Uchida, M., Gershman, S.J., and Uchida, N. (2020). A Unified Framework for Dopamine Signals across Timescales. Cell *183*, 1600–1616.e25.
34. Shine, J.M., Müller, E.J., Munn, B., Cabral, J., Moran, R.J., and Breakspear, M. (2021). Computational models link cellular mechanisms of neuromodulation to large-scale neural dynamics. Nat. Neurosci. *24*, 765–776.
35. Záborszky, L., Gombkoto, P., Varsanyi, P., Gielow, M.R., Poe, G., Role, L.W., Ananth, M., Rajebhosale, P., Talmage, D.A., Hasselmo, M.E., et al. (2018). Specific Basal Forebrain–Cortical Cholinergic Circuits Coordinate Cognitive Operations. J. Neurosci. *38*, 9446–9458.
36. Shine, J.M. (2019). Neuromodulatory Influences on Integration and Segregation in the Brain. Trends Cogn. Sci. *23*, 572–583.
37. Fuxreiter, M., and Warshel, A. (1998). Origin of the Catalytic Power of Acetylcholinesterase: Computer Simulation Studies. J. Am. Chem. Soc. *120*, 183–194.
38. Disney, A.A., and Higley, M.J. (2020). Diverse Spatiotemporal Scales of Cholinergic Signaling in the Neocortex. J. Neurosci. *40*, 720–725.
39. Sarter, M., and Lustig, C. (2020). Forebrain Cholinergic Signaling: Wired and Phasic, Not Tonic, and Causing Behavior. J. Neurosci. *40*, 712–719.
40. Jing, M., Zhang, P., Wang, G., Feng, J., Mesik, L., Zeng, J., Jiang, H., Wang, S., Looby, J.C., Guagliardo, N.A., et al. (2018). A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. Nat. Biotechnol. *36*, 726–737.
41. Vertes, R.P. (1988). Brainstem afferents to the basal forebrain in the rat. Neuroscience *24*, 907–935.
42. España, R.A., and Berridge, C.W. (2006). Organization of noradrenergic efferents to arousal-related basal forebrain structures. J. Comp. Neurol. *496*, 668–683.
43. Joshi, S., and Gold, J.I. (2019). Pupil Size as a Window on Neural Substrates of Cognition. https://doi.org/10.31234/osf.io/dvsme.
44. Joshi, S., Li, Y., Kalwani, R.M., and Gold, J.I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron *89*, 221–234.
45. Cazettes, F., Reato, D., Morais, J.P., Renart, A., and Mainen, Z.F. (2021). Phasic Activation of Dorsal Raphe Serotonergic Neurons Increases Pupil Size. Curr. Biol. *31*, 192–197.e4.
46. Keehn, B., Kadlaskar, G., Bergmann, S., McNally Keehn, R., and Francis, A. (2021). Attentional Disengagement and the Locus Coeruleus – Norepinephrine System in Children With Autism Spectrum Disorder. Front. Integr. Neurosci. *15*, 716447.
47. Kim, Y., Kadlaskar, G., Keehn, R.M., and Keehn, B. (2022). Measures of tonic and phasic activity of the locus coeruleus—norepinephrine system in children with autism spectrum disorder: An event-related potential and pupillometry study. Autism Res. *15*, 2250–2264.
48. Yu, S., Ghin, F., Mückschel, M., Ziemssen, T., Stock, A.-K., and Beste, C. (2022). A role of the norepinephrine system or effort in the interplay of different facets of inhibitory control. Neuropsychologia *166*, 108143.
49. Frost, S., Robinson, L., Rowe, C.C., Ames, D., Masters, C.L., Taddei, K., Rainey-Smith, S.R., Martins, R.N., and Kanagasingam, Y. (2017). Evaluation of Cholinergic Deficiency in Preclinical Alzheimer's Disease Using Pupillometry. J. Ophthalmol. *2017*, 7935406.
50. Mesulam, M.-M., Mufson, E.J., Levey, A.I., and Wainer, B.H. (1983). Cholinergic innervation of cortex by the basal forebrain: Cytochemistry and cortical connections of the septal area, diagonal band nuclei, nucleus basalis (Substantia innominata), and hypothalamus in the rhesus monkey. J. Comp. Neurol. *214*, 170–197.
51. Obermayer, J., Luchicchi, A., Heistek, T.S., De Kloet, S.F., Terra, H., Bruinsma, B., Mnie-Filali, O., Kortleven, C., Galakhova, A.A., Khalil, A.J., et al. (2019). Prefrontal cortical ChAT-VIP interneurons provide local excitation by cholinergic synaptic transmission and control attention. Nat. Commun. *10*, 5280.

52. Granger, A.J., Wang, W., Robertson, K., El-Rifai, M., Zanello, A.F., Bistrong, K., Saunders, A., Chow, B.W., Nuñez, V., Turrero García, M., et al. (2020). Cortical ChAT+ neurons co-transmit acetylcholine and GABA in a target- and brain-region-specific manner. Elife *9*, e57749.

53. Bayraktar, T., Staiger, J.F., Acsady, L., Cozzari, C., Freund, T.F., and Zilles, K. (1997). Co-localization of vasoactive intestinal polypeptide, g-aminobutyric acid and choline acetyltransferase in neocortical interneurons of the adult rat. Brain Res. *757*, 209–217.

54. Von Engelhardt, J., Eliava, M., Meyer, A.H., Rozov, A., and Monyer, H. (2007). Functional Characterization of Intrinsic Cholinergic Interneurons in the Cortex. J. Neurosci. *27*, 5633–5642.

55. Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., et al. (2017). Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data.

56. Xu, M., Chung, S., Zhang, S., Zhong, P., Ma, C., Chang, W.-C., Weissbourd, B., Sakai, N., Luo, L., Nishino, S., and Dan, Y. (2015). Basal forebrain circuit for sleep-wake control. Nat. Neurosci. *18*, 1641–1647.

57. Lee, M.G., Hassani, O.K., Alonso, A., and Jones, B.E. (2005). Cholinergic Basal Forebrain Neurons Burst with Theta during Waking and Paradoxical Sleep. J. Neurosci. *25*, 4365–4369.

58. Yaksi, E., and Friedrich, R.W. (2006). Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging. Nat. Methods *3*, 377–383.

59. Evans, M.H., Petersen, R.S., and Humphries, M.D. (2019). On the Use of Calcium Deconvolution Algorithms in Practical Contexts. bioRxiv. https://doi.org/10.1101/871137.

60. Friedrich, J., Zhou, P., and Paninski, L. (2017). Fast online deconvolution of calcium imaging data. PLoS Comput. Biol. *13*, e1005423.

61. Holekamp, T.F., Turaga, D., and Holy, T.E. (2008). Fast Three-Dimensional Fluorescence Imaging of Activity in Neural Populations by Objective-Coupled Planar Illumination Microscopy. Neuron *57*, 661–672.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat anti ChAT	Chemicon	Cat. AB144P
AF546 anti Goat	Invitrogen	Cat. A-21085
Chicken anti GFP	abcam	Cat. Ab13970
AF488 anti Chicken	Invitrogen	Cat. A11039
Bacterial and virus strains		
AAV-hSyn-ACh3.0	Vigene Biosciences	N/A
AAV-hSyn-rACh1.0	WZ Biosciences	N/A
AAV9-hSyn-ACh3.0mut	WZ Biosciences	N/A
AAV-*syn*-FLEX-jGCaMP8s	addgene	N/A
Chemicals, peptides, and recombinant proteins		
Acetylcholine chloride	Sigma Aldrich	N/A
Norepinephrine bitartrate	Sigma Aldrich	N/A
Deposited data		
Data		https://dandiarchive.org/dandiset/001176
Original code		https://zenodo.org/records/13362427
Experimental models: Organisms/strains		
C57BL/6J mice		N/A
ChAT-IRES-Cre mice		N/A
ChAT-IRES-Cre (Δneo) mice		N/A
Software and algorithms		
Python		N/A
Datajoint		N/A
Docker		N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This study used a total of 34 mice aged 1.5 to 12 months (n = 13 male, n = 21 female). Genotypes included wild type C57Bl/6 mice (or various other negative genotypes on a C57Bl/6 background) (n = 29), ChAT-cre (n = 1), and Δ*neo* ChAT-cre (n = 4). All animals were singly housed in the animal care facility at Baylor College of Medicine (BCM), and before participating in this study mice had not undergone any previous procedures. All experimental procedures were carried out in accordance with the ethical guidelines of the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee (IACUC) of BCM. Because we did not observe a substantial difference in the magnitude of run- or dilation-triggered GRAB-ACh in male vs. female mice, we do not expect sex to have a significant effect on the results of this study. Additionally, when comparisons between animals were conducted, we included roughly an equal ratio of male to female mice in each experimental cohort (or, where this was not possible, a roughly equal ratio of scans from male vs. female mice) in order to equally distribute any potential sex differences across groups.

METHOD DETAILS

Animals, surgery, and injection of viral vectors

All procedures were carried out in accordance with the ethical guidelines of the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee (IACUC) of Baylor College of Medicine. This study used a total of 34 mice aged 1.5 to 12 months (n = 13 male, n = 21 female).

Surgical procedures for spontaneous state-change experiments, ACh injection experiments, and hemodynamic control experiments

Under isoflurane anesthesia, wild type C57Bl/6 mice (or various other negative genotypes on a C57Bl/6 background) were implanted with cranial windows of the following size and placement: 4 mm over visual cortex (n = 18), 3 mm over somatosensory cortex (n = 3),

3mm over motor cortex ($n = 2$), 4 mm over motor cortex ($n = 3$), or 5 mm spanning visual and somatosensory cortices ($n = 3$). The head was shaved over the dorsal portion of the skull and bupivacaine 0.25% (0.1 mL) and ketoprofen (5 mg/kg) were administered subcutaneously. The skin was pulled back and glued in place using Vetbond (3M). A headbar (steel washer with outer diameter 0.438″, inner diameter 0.314″, thickness 0.016″) was affixed to the skull using dental cement (C&B Metabond). A micromotor rotary drill (Foredom) fitted with a 0.6 mm tungsten carbide bur was used to thin the skull over the center of the chosen cortical area, which was determined by stereotactic coordinates (VIS: 2.8 mm lateral, 1 mm anterior lambda; VIS & SS: 2.8 mm lateral, 2.25 mm posterior bregma; SS: 3.25 mm lateral, 0.3 mm posterior bregma; MO: 1.5 mm lateral, 1 mm anterior bregma), and a circle of skull was removed. A borosilicate glass pipette (World Precision Glass) was pulled to a tip width of 40μm and attached to a syringe which was used to inject 500-2000nL of AAV-hSyn-ACh4.3 at a depth of 200-700μm below the surface using a Micro4 MicroSyringe pump controller. For hemodynamic control experiments, mice received either a single 1300 nL injection of a 4:1 mix of rACh and GRAB-ACh-mut ($n = 1$), or 250 nL each of GRAB-ACh and GRAB-ACh-mut injected in spatially segregated locations in the window ($n = 3$). The dura was removed, and a #1 glass coverslip (thickness 0.15 ± 0.02 mm) was then glued in place using super glue (Loctite). The animal was allowed to recover in its home cage on a heated disk until sternal recumbency was regained, and was monitored for the next 3 days for signs of pain or distress.

Surgical procedures for dual axon-ACh imaging experiments

ChAT-cre ($n = 1$) or Δ*neo* ChAT-cre ($n = 4$) mice first received an injection into the horizontal diagonal band (HDB) of the basal forebrain after anesthesia induction by i.p. Injection of ketamine and xylazine. HDB was selected due to its preferential cholinergic projections to the visual cortex.[26–28] The head was shaved and skin pulled back after an incision was made. A burr hole was drilled and, micropipette fitted with a 20μm glass pipette was positioned to AP: 0.13, ML: 1.1, DV: −5.6, and 1600nL of AAV-*syn*-FLEX-jGCaMP8s was injected over 160 s. The skin was sutured back in place using braided absorbable sutures (McKesson) and the animal allowed to recover. After 1 week, a headbar and 4mm cranial window was implanted over the visual cortex using the above procedures, into which 1000-1800nL of AAV-hSyn-rACh1.0 was injected at a depth of 300μm below the surface.

In vivo pharmacological experiments

In 4 mice under general isoflurane anesthesia, the glass coverslip previously implanted over the visual cortex was removed and replaced with a coverslip of the same size but with a small hole near the edge. The mouse was head-fixed under the microscope where it remained anesthetized for the duration of the experiment. The following agents were diluted with cortex buffer and Alexa 568 dye to achieve specific molecular concentrations: acetylcholine chloride (crystals/powder, Sigma-Aldrich), norepinephrine bitartrate (crystalline, Sigma-Aldrich). A borosilicate patching pipette was fitted to a syringe and backfilled with the mixed pharmacological agent and inserted into the cortex through the hole in the coverslip. 200 mBar pressure was applied to the syringe, which was released with a button press.

Histology

After the completion of two-photon imaging, mice were deeply anesthetized with isoflurane and transcardially perfused with 4% paraformaldehyde (PFA). Brains were removed from skulls and stayed in 4% PFA overnight at 4°C. 100-μm-thick coronal sections were cut with a vibratome (Leica 1000S) in 0.1 M PBS. Unspecific antigens were blocked by incubating selected sections for 1 h in 10% normal donkey serum (NDS, Sigma, Cat# D9663) and 0.5% Triton X-100 in 0.1 M PBS. Sections were selected that contained the HDB injection sites and Flex-jGCaMP8s labeled axons in VIS. Following the blocking step, the selected sections were incubated in a 1:300 dilution of goat polyclonal antibody against ChAT (Millipore Cat# AB144P, RRID: AB_2079751), 2% NDS and 0.5% Triton X-100 in 0.1 M PBS overnight at 4°C. On the second day, the sections were washed with 0.01 M PBS three times for 10 min each and then incubated in a 1:100 dilution of a rabbit-*anti*-goat antibody that was directly conjugated to Alexa Fluor 546 (Invitrogen, Cat# A-21085, RRID: AB_2535742), 2% NDS and 0.5% Triton X-100 in 0.1 M PBS for 2 h. The sections were then washed with 0.01 M PBS and mounted on slides using VECTASHIELD antifade mounting medium containing DAPI (VectorLabs, Cat# H-1200-10).

Locomotion and pupillometry

For all experiments, the mouse was positioned atop a 9″-diameter x 3″-wide cylindrical styrofoam treadmill and head-fixed using a steel headbar clamped onto the surgically implanted washer. The surface of the treadmill was covered in duct tape to provide extra traction. The mouse's body was free to move with 2 degrees of freedom (forward and backward). An optical encoder (McMaster-Carr) was used to keep track of treadmill speed. Treadmill traces were processed with a 0.5 s median filter. Running periods were defined as running at a speed of least 1 cm/s for at least 1 s. Periods less than 3 s apart were grouped together into a single period. Periods of quiescence were determined by removing all running periods - 3 s.

A Dalsa camera mounted with a 55mm telecentric lens (OEM) and a 1″ hot mirror (Thorlabs) was used to capture a video of the mouse's eye at 20 fps. As the experiment was conducted in darkness with no visual stimulus, the eye was illuminated from an outside source (either a computer monitor or a UV light) to prevent total pupil dilation and allow normal, spontaneous fluctuations in pupil size to be measured.

A DeepLabCut model was used to identify pupil edges from the eye videos, from which the circumference of the pupil was interpolated. The radius of the pupil (in pixels) was then determined, and this trace was 1hz lowpass filtered. Radius traces were inspected by eye to determine quality of tracking, and movies which had poor tracking were thrown out. To convert radius in pixels to radius in

millimeters, the corners of the eyelid were detected using DeepLabCut and the median distance was assumed to be 2.757mm (which was determined by our measurements to be the universal mouse eyelid width). This was used to create a conversion factor whereby pixels could be converted to millimeters. Dilation (or constriction) events where then defined as periods of increasing (decreasing) pupil radius lasting greater than 1 s and with an average speed greater than 0.02 mm/s. Blinks or frames where the pupil was not able to be detected by DeepLabCut were considered NaNs, and any event comprised of more than 15% NaNs was dropped.

Imaging

All scans were recorded using a two-photon fast resonant scanning system (ThorLabs). Frame rates varied from 5 to 60fps. Excitation wavelengths were either 920nm (GRAB-ACh scans) or 1000nm (simultaneous GCaMP and rACh scans) with a 25× objective (Nikon). Output power was kept around 15-40mW. The green emission filter was 520-550nm and the red emission filter was 620-670nm. ACh injection scans were imaged at a frame rate of 58 fps with an FOV size around 200 × 200 μm. Simultaneous GCaMP and rACh scans were imaged at a frame rate of 30 fps with an FOV size around 100 × 100 μm. All other GRAB-ACh scans were imaged at a frame rate ranging from 3 to 58 fps with an FOV size ranging from 230 × 230 μm to 3000 × 600 μm.

Motion correction and selection of ROIs

All scans were corrected for motion in the X and Y directions via a two-step algorithm. Large scale motion was calculated by taking the average of the movie and computing phase correlation between this template and each frame of the movie. After correcting for large scale motion, smaller scale motion was corrected by taking the average of a 2,000-frame window and phase correlating frames within this window to the local template. Windows were selected to have 500 overlapping frames, and a linearly weighted average was taken for the displacement in the overlapping region.

Scans were discarded which had poor motion correction (defined as root-mean-squared of the x- and y-shifts produced from motion correction above 4 μm) or which contained large, unexplained jumps in fluorescence (at least 20x the average derivative). Three scans were discarded in which, despite adequate motion correction, it was assumed that the recorded axons moved out of the recording plane during motion due to a complete absence of detected spikes during running periods.

Fluorescence traces for each scan were determined by taking the average fluorescence of all pixels within a certain ROI (“mask”). For scans which recorded GRAB-ACh or rACh sensors, masks were determined by taking pixels which had average brightness values < 99.9th percentile and >(99.9th percentile/2) or >(99th percentile/2), respectively, after scan borders were removed (see Figure S5). The more inclusive lower-bound threshold used for the rACh scans was selected due to overall dimmer expression of the red sensor versus the green sensor. Each mask was then eroded using a 7x7-pixel kernel to provide smoothness. For axon scans, because responses to running were mostly homogeneous across all axon segments within a scan (Figure S6), a single mask which covered all visible axons was used for the green (GCaMP) channel. This mask was determined by calculating the overlap between a manually drawn mask over visible axons and a brightness mask (i.e., pixels with an average brightness value greater than (99th percentile/2) after removing borders). For scans where ACh was injected, a circular mask with 50μm radius was created over the tip of the injection pipette, which was manually defined.

For distance-binned peri-spike traces, we separately analyzed areas of the FOV containing axons and non-overlapping areas without axons (i.e., containing only rACh fluorescence) by dividing the FOV into six ROIs. In each scan, the ROI which contained axons (“distance bin 0”) was determined using the overlap of a manually drawn mask and a brightness threshold Figures 5A and S5C, navy blue pixels). The remaining portion of the FOV was divided pixel-wise into five ROIs (distance bins 1–5; determined based on the cumulative distribution of distances across all scans in Figure S1) based on each pixel’s distance from the nearest axon-containing ROI. We analyzed ROIs with and without axons separately in order to be conservative about eliminating effects of bleed-through GCaMP fluorescence. Although it is unlikely that rACh is expressed on the axonal membrane of these projections from the BF since the sensor is injected locally in the cortex, our method of separating axon-containing ROIs would also eliminate the effects of any axonal rACh expression.

Characterization of and control for hemodynamic artifact

Hemodynamic artifacts are always a concern when recording bulk fluorescence with a sensor such as GRAB-ACh (although much less of a concern than with one-photon methods). With GRAB-ACh, we have found such an artifact in the form of a brief decrease (“dip”) in fluorescence which specifically occurs at the onset of locomotion. This dip is observed both with non-responding mutated versions of the sensor and also with GFP. In our characterization of this artifact, we’ve discovered that this dip is more prominent when recording at deeper cortical depths than at superficial depths (Figures S7A and S7B), as expected for fluorescence signals collected through larger (deeper) volumes of tissue. To be conservative, we excluded GRAB-ACh scans from our analysis which were recorded at or deeper than 100μm in VIS (a small number of scans recorded between 100 and 200μm in SS and MO were left in the analysis due to the limited number of scans in those areas).

To further characterize this artifact, we performed both 2-channel and 2-field (i.e., 2-ROI) controls with GRAB-ACh-mut and rACh or GRAB-ACh (Figures S7C and S7E). In the 2-channel recordings, although a run-related artifact was found in GRAB-ACh-mut, no such dip was found in the rACh signal (Figure S7D). In this case, correcting the rACh trace with the mut trace did not result in

qualitative or quantitative differences in the rACh signal. In the two-field recordings, a run-related artifact was observed in both fields (Figure S7F). While correcting the GRAB-ACh trace with the mut trace did remove the artifact from run onset, it again did not result in qualitative or quantitative differences in the GRAB-ACh signal.

Both of the above correction methods are potentially flawed. When imaging two simultaneous fields, the differing vasculature patterns between the two sites could potentially lead to differences in artifact characteristics leading to an incorrect correction. When imaging two channels, the different fluorophores (red vs. green) could similarly contribute to differences in artifact characteristics. After extensive characterization, we believe that for the GRAB-ACh sensor, the signal to noise appears to be relatively high - we did not observe a prominent hemodynamic artifact with rACh, and the two-field correction with GRAB-ACh-mut did not substantially alter the characteristics of the GRAB-ACh signal. Thus, our approach is to present the minimally processed fluorescence after removing deeper scans that might be more impacted by hemodynamic changes. However, caution should be used with less sensitive sensors with signals that may be dominated by vascular artifacts.

Peri-dilation, peri-run, and peri-spike traces

Fluorescence traces were first preprocessed as follows. The first 20 s of each trace was thrown out in order to account for stabilization of the laser. Traces were then detrended to correct for photobleaching by subtracting a 4^{th}-order polynomial, 1hz lowpass filtered (except for Figure 5B, which used a 2hz lowpass filter), and converted to $\Delta F/F$. Baseline F was calculated using a 20-min non-causal rolling 5^{th}-percentile and $\Delta F/F$ was calculated as (fluorescence - baseline)/baseline. Our scanning acquisition software attempts to center noise levels at 0, which means that it is possible for raw fluorescence values to fall below 0 (i.e., contain negative values) for scans with very low overall fluorescence. Therefore, traces which had negative baseline values were discarded in order to avoid dividing by a negative baseline. Furthermore, dim scans which had a low baseline but also a high range of values resulted in an abnormally high $\Delta F/F$ values; therefore, scans that had a ratio of (range of fluorescence values): (5^{th} percentile of fluorescence values) greater than 5 were considered outliers and discarded. Deconvolved traces in Figure 6 were preprocessed in this order: detrending, deconvolution (see STAR Methods in section below), filtering, $\Delta F/F$ normalization.

The times at which dilation, run onset, and run offset events occurred were determined using the methods in the above section ("Locomotion and pupillometry"). Peri-event traces were calculated by subtracting the average pre-event fluorescence value ("baseline") and multiplying by 100 to convert to a percentage. Peri-dilation events containing NaNs in the radius trace were excluded. All peri-event traces were averaged together and standard error was found by dividing the standard deviation of all traces by the square root of the number of traces.

To determine spike times for peri-spike traces, the raw fluorescence calculated from the axon mask described above ("Motion correction and selection of ROIs") was deconvolved using a non-negative matrix factorization algorithm.[55] For each trace, a threshold was set at the 96^{th} percentile of the deconvolved fluorescence values and every sample point at or above this threshold was defined as a spike time. This threshold level was selected in order to achieve spiking rates which matched biologically-plausible firing rates of the cholinergic basal forebrain (approximately 4 spikes/sec overall[56] and approximately 7 spikes/sec during active periods[57]).

Coherence and cross-correlation

Coherence between pupil radius and ACh sensor activity or GCaMP axon activity was performed on unfiltered radius and fluorescence traces using a 1-min (1,200-sample) long Discrete Prolate Spheroidal Sequences window with overlap of 1,100 and half-bandwidth of 4. Fluorescence traces were detrended and radius traces had NaNs interpolated before calculating coherence.

Cross-correlations were performed by first normalizing pupil radius and fluorescence traces. Unfiltered and NaN-interpolated pupil radius traces were normalized by subtracting the mean and dividing by the standard deviation times the number of samples in the trace. Unfiltered and detrended fluorescence traces were normalized by subtracting the mean and dividing by the standard deviation of the trace. Cross-correlations were then calculated on the normalized traces, and peaks/lags were determined using the middle $+/-$ 2 s of the cross-correlations.

GRAB-ACh deconvolution

Fluorophore impulse response

Recent empirical work had characterized the GRAB-ACh sensor kinetics[21]; we first used those results to model the fluorophore's impulse response (or convolution/deconvolution kernel) k. The exponential fall time had been directly determined to be 580ms, by measurement of fluorescence in the presence of a competitive agonist. The inverse rise-time of the fluorescence response to an ACh puff increases logarithmically with molarity, with a constant $3.12/(\mu M \times s)$. However, as the convolution kernel is defined as the response to an instantaneous pulse of infinite concentration, and moreover because these are association kinetics (rather than kinetics of conformation state change in the fluorophore itself) by extrapolation the kernel's rise time should be effectively zero; we therefore modeled the kernel as a single falling exponential with time constant 580ms.

Convolutional model

We then assumed (cf. Yaksi and Friedrich, 2006[58]) that the fluorescence signal is generated by convolution of a time-varying concentration $[ACh](t)$ with the fluorophore's impulse response k, so that the GRAB-ACh fluorescence trace $F(t)$ is:

$$F(t) = \int_{\tau} k(t-\tau)[ACh](t)d\tau$$

Roughly, the instantaneous concentration $[ACh]$ at time τ will cause a characteristic fluorescence response $k(t-\tau)$ – which lasts for some time after fluorophores are activated. Contributions from fluorophores so activated at various times in the past (i.e., $t-\tau$) are then integrated to generate the fluorescence at time t. Note that this model does not account for nonlinear effects, in particular saturation.

Deconvolution

Deconvolution methods then attempt to invert the process of convolution, recovering the source $[ACh]$ from the fluorescence signal and the impulse response of the fluorophore. The simplest deconvolution method divides the Fourier transform (denoted $\mathcal{F}$, with inverse transform $\mathcal{F}^{-1}$) of the fluorescence signal by that of the kernel:

$$[ACh] \propto \mathcal{F}^{-1}\left(\frac{\mathcal{F}(F(t))}{\mathcal{F}(k)}\right)$$

However, we found this approach led to estimates of $[ACh]$ which – at least at the level of individual scans – were so noisy as to be uninterpretable.

More advanced deconvolution methods differ largely by their assumptions on source and noise, yielding a denoised estimate of a conditioned source. Several modern methods (see also the review[59]), for instance, deconvolve latent voltage spikes from calcium fluorescence.[60] In our study, however, we cannot assume *a priori* that acetylcholine concentration $[ACh]$ is sparse or spikelike. A classical approach – Wiener deconvolution (which had also previously been applied, with some modification, in the calcium context[61]) – yielded intelligibly denoised single-trial $[ACh]$ estimates without further source constraints, making it a suitable method for our purposes.

Wiener deconvolution modifies the division method above by adding a regularization parameter to the denominator:

$$[ACh] \propto \mathcal{F}^{-1}\left(\frac{\mathcal{F}(F(t))}{\mathcal{F}(k)+\lambda^2}\right)$$

We estimated this parameter λ based on empirical considerations: As it increases, the reconstructed signal first shows removal of higher-frequency noise, but then shows a flattening of the variation of interest at lower frequencies – eventually approaching a constant value (See methods video S1). To select an appropriate degree of smoothing for each individual scan, we first standardized the signal, then varied lambda until the absolute reconstruction error was 10% of its maximum (i.e., 10% of when the signal flattened completely). Although this 'optimal' value of lambda was slightly larger computed over entire scans vs. stationary-only periods ($p < 10^{-10}$; Wilcoxon test), the size of this difference was fairly small (mean: 11.3%).

Axonal to ACh predictive modeling

The axonal to ACh predictive model related the axonal fluorescence (F) to the simultaneously recorded rACh signal. First, a sigmoidal function calculated the maximal axonal output, $S(F)$,

$$S(F) = \frac{1}{1+e^{\left(-\frac{F-\theta}{\sigma}\right)}},$$

where θ horizontally shifts the sigmoid (e.g., high $\theta \rightarrow$ smaller output S for given activity F) and σ is the rate of growth of the sigmoid from baseline to threshold. Next, we simulated the temporal release of ACh following an exponential rise and decay function by temporally convolving $S(F)$ with the release function $R(t)$,

$$R(t) = \frac{A\left(1-e^{-\frac{t}{\tau_r}}\right)\times e^{-\frac{t}{\tau_d}}}{\sum_{t=t_{min}}^{t=t_{max}}\left(1-e^{-\frac{t}{\tau_r}}\right)\times e^{-\frac{t}{\tau_d}}},$$

where τ_r and τ_d are the release rise and decay times, A is a scaling parameter, and the denominator is a normalization factor. Finally, the predicted ACh is convolved with the sensor dynamics $e^{-\frac{t}{\tau_{off}}}$ (normalized to unit area) with $\tau_{off} = 580$ ms (estimated from GRAB-ACh kinetics,[21] due to rACh kinetics being yet unpublished). The optimization swept across the 5 parameters $(\theta, \sigma, A, \tau_r, \tau_d)$ using an interior-point algorithm (MATLAB fmincon) over 150 iterations to minimize the RMSE between the predicted and empirically recorded GRAB-ACh signal.

Behavioral to ACh predictive modeling

We predicted GRAB-ACh fluorescence from behavioral signals – running speed, pupil size and temporal derivatives – using a long short-term memory (LSTM) recurrent neural network. Briefly, the RNN consisted of an input layer, an LSTM layer (with a working memory of the previous ~8.5 s), a fully connected 256-node recurrent neural network, a 10% dropout layer, and a fully connected output layer. The model was trained on ~ 2-min epochs across 90% of the recordings and validated on unseen held-out recordings (10% of total recordings). We optimized the parameters of the network through stochastic gradient descent using the Adam optimizer over 150 iterations, minimizing the RMSE between predicted and observed GRAB-ACh. The predictive model was trained using running speed, pupil size and their temporal derivatives (pupil + speed), pupil size and its temporal derivative (pupil), and running speed and its temporal derivative (speed). Model fits were assessed by calculating the Pearson correlation between predicted and observed GRAB-ACh in the held-out validation dataset.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was performed using custom Python code, and particular statistical details can be found in figure legends. In all figures utilizing statistical comparisons, N's were defined as number of running periods, dilation periods, pharmacological injections, or spikes. Comparisons were visualized with peri-event time histograms, where baseline was defined as the average activity in the pre-event window and this baseline was subtracted from the entire activity trace in the window (i.e., +/− X seconds before/after event). Mean traces were plotted, with error bars representing +/− SEM. Statistical comparisons were conducted using Kruskal-Wallis tests to compare medians, and significance was defined as $p < 0.05$.