Multitimescale Computation by Astrocytes

Authors: Chang Li[1*], Lulu Gong[2], Chenghui Song[1], ShiNung Ching[3], Lucas-Pozzo Miller[4], Wei Li[1,4*]

Affiliations:

1. Department of Neurobiology, University of Alabama at Birmingham, Birmingham, AL, United States

2. Department of Biomedical Engineering Wu Tsai Institute Center for Neurocomputation and Machine Learning Yale University, New Haven, CT, United States

3. Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, United States

4. Department of Pediatrics, College of Human Medicine, Michigan State University, Grand Rapids, MI, United States

*Corresponding authors; email: lic@uab.edu; liwei24@msu.edu

Author Contributions: CL: conceptualization, investigation, formal analysis, visualization, project administration, writing – original draft, and writing – review & editing. LG: methodology, formal analysis, software, validation, visualization, and writing – review & editing. CS: software, data curation, formal analysis, and writing – review & editing. SC: methodology and writing – review & editing. LP-M: conceptualization, funding acquisition, resources, supervision, and writing – review & editing. WL: conceptualization, funding acquisition, resources, supervision, project administration, and writing – review & editing.

Acknowledgements: This work was supported by the National Institutes of Health grants R01-NS121542, R21-NS108508, and R21-NS120315. We thank Dr. Vladimir Parpura for helpful comments. We are grateful to Yijian Zhang, Likhitha Polepalli, and Jenny Shen for maintaining the mouse colony. We also thank lab members Julia Lopes Goncalez, César Acevedo-Triana, Xin Xu, Destynie Medeiros, Suraj Cherian, and Akash Saxena for their assistance and support.

Abstract

The basic computational unit of the brain has long been defined as the neuron. However, mounting evidence suggests that other cells, especially astrocytes, also perform computation. Here we demonstrate that cerebellar astrocytes decompose norepinephrine input into slow and fast calcium activities through differential adrenergic receptor engagement. During reward learning in mice, slow and fast activities selectively target distinct synaptic pathways. Causal manipulations reveal that slow α1-adrenergic signals maintain behavioral states and coordinate transitions, while fast α2-adrenergic signals govern event-triggered responses and reinforcement learning. Remarkably, an actor-critic neural network trained on a similar sequence task spontaneously recapitulates these

multitimescale dynamics, suggesting astrocytes implement critic-like computations that evaluate states and modulate neuronal learning. These findings establish astrocytes as multilevel processors that transform univariate neuromodulatory inputs into multivariate, pathway-specific circuit control operating in parallel with neuronal processing.

Introduction

Neurons have traditionally been considered the brain's fundamental computational units[1-3]. However, astrocytes exhibit signaling spanning orders of magnitude in time[4,5], contact thousands of synapses[6-8], and receive neuromodulatory inputs through diverse receptors[9,10], suggesting they may also perform computation.

Traditional characterizations of astrocytic calcium signals as slow, uniform events[11-13] have supported modulatory roles adjusting circuit tone[14,15]. Recent studies identified faster transients[16-18], but whether these signals transform neuromodulatory inputs into functionally distinct outputs that causally regulate behavior remains uncharacterized[19-21]. The cerebellar cortex provides an ideal model: climbing fibers (CF) carry instructive signals[22,23] and parallel fibers (PF) convey contextual information[24,25], converging onto Purkinje cells, the sole output of the cerebellar cortex[26,27]. Bergmann glia (BG), the neighboring astrocytes of Purkinje cells, interact with both synapses[28,29], positioning them to differentially regulate these pathways.

Here we show BG decompose norepinephrine into slow and fast calcium activities through differential α-adrenergic receptors, with each selectively modulating distinct synaptic pathways and controlling separable behavioral functions. Causal manipulations confirm that disrupting each activity mode impairs distinct aspects of behavior. An actor-critic network trained on a similar task spontaneously develops these multitimescale dynamics, suggesting astrocytes implement computations that operate in parallel with neurons to coordinate behavior.

Results

Cerebellar Astrocyte Ca^{2+} Dynamics During Reward-Guided Behavior

To examine BG Ca^{2+} activities during behavior, we trained mice on a spatial sequence task requiring movement from a trigger zone to a reward zone (Fig. 1a). Sequence completion triggered sucrose delivery after 2 s, followed by a 4 s reward window. Mice trained for 30 min per day for 15 days, with Day 1 serving as a no-reward baseline. Infrared sensors detected trigger and reward zone entry/exit, reward delivery, and licking, enabling precise alignment of behavior and BG Ca^{2+} activities.

We recorded BG Ca^{2+} using fiber photometry[30,31] in Aldh1l1-CreERT2;GCaMP6f mice[32,33] targeting left and right simplex and lobule III (Fig. 1b; Supplemental Fig. 1). Animals progressively completed more trials, transitioning faster between trigger and reward zones, and restricted licking to the reward window, demonstrating emergence of goal-directed behavior (Supplemental Fig. 2).

Example recordings show progression across training: Day 1 exhibited variable behavior with minimal signals, while Day 15 showed robust, temporally structured behavior and Ca^{2+} activities (Fig. 1c). Raw traces contained both slow ramping and brief fast transients; filtered traces highlighted slower dynamics. Heatmaps aligned to licking onset and trigger zone exit (Fig. 1d) show that slow ramping emerged during licking with training, while fast transients became prominent at trigger zone exit.

Slow BG Ca^{2+} Activities Encode Behavioral State Transitions

We classified slow BG Ca^{2+} dynamics during licking into three polarity-based types—trough, state-change, and peak—based on responses at lick onset in trained animals (Fig. 1e-g). Trough signals declined before licking, reached minimum at onset, and peaked at trigger zone entry, suggesting reward suppression. State-change signals transitioned from positive to negative across the lick window and reversed polarity at reward zone exit, consistent with shifts between reward-seeking and engagement. Peak signals ramped up before licking, peaked at onset, and declined toward reward zone exit, consistent with reward activation.

All three signal types appeared across cerebellar regions without clear anatomical bias (Supplemental Fig. 3). Notably, summing average peak and trough traces produced a composite resembling the state-change trace, suggesting state-change dynamics arise from integration of opposing responses across regions.

Re-aligning traces to reward onset and performing slope analysis (Fig. 1h) revealed consistent temporal trends across pre-reward (−4 to 0 s), reward (0 to 4 s), and post-reward (4 to 8 s) periods. Trough signals were modulated before and after reward, while peak and state-change signals shifted during and after delivery. Comparing rewarded versus unrewarded spontaneous licking showed reward presence strongly influenced dynamics, with differences emerging before reward onset, suggesting preparatory components (Fig. 1i-k).

Fast BG Ca^{2+} Transients Show Hemisphere-Specific Tuning

Fast BG Ca^{2+} transients exhibited spatial distribution and learning modulation. In left simplex, reward zone entry evoked strong responses on Day 1 that diminished with training (Fig. 1l). Conversely, trigger zone exit—preceding reward approach—elicited signals that increased with training (Fig. 1m). No response occurred at trigger zone entry (tone onset), suggesting auditory cues alone did not drive BG activity (Fig. 1n).

In right simplex, fast signals aligned to trigger zone entry (tone onset) and gradually increased with training (Fig. 1q), while responses to reward zone entry and trigger zone exit remained flat (Fig. 1o-p). Raw traces revealed prominent fast events in hemispheres but not midline lobule III (Supplemental Fig. 4).

NE-Dependent Modulation of BG Ca^{2+} Activities

We tested whether norepinephrine (NE) modulates these activities during behavior. NE is a key neuromodulator released from locus coeruleus projections to the cerebellum[34-36] implicated in

cerebellar learning[37-39]. Using GRAB_NE sensors[40] (Fig. 2a), we found NE transients followed the same timing and trends as the three slow BG signal types (Fig. 2b-d), indicating slow BG Ca^{2+} tracks NE fluctuations. Re-alignment to reward onset showed NE displayed similar patterns but with sharper changes during reward (Fig. 2e).

Fast NE dynamics aligned to trigger zone exit (left hemisphere, Fig. 2g) and trigger zone entry (right hemisphere, Fig. 2h) showed training-dependent activity similar to astrocytic fast transients. Reward zone entry showed minimal NE response (Fig. 2f), differing from Ca^{2+} responses. These findings suggest NE acts as a reinforcement signal during cue presentation and reward approach.

α1-ARs and α2-ARs Mediate Distinct BG Ca^{2+} Dynamics

We combined fiber photometry with local drug infusion to test adrenergic receptor contributions (Fig. 2i). Yohimbine (α2-AR antagonist) decreased fast BG Ca^{2+} amplitude at trigger zone exit (left hemisphere) and entry (right hemisphere, Fig. 2j-l), confirming α2-ARs mediate fast transients. Yohimbine did not affect slow activities (Supplemental Fig. 5).

Prazosin (α1-AR antagonist) decreased slow signals across events (Fig. 2m-p), reducing peak amplitude and shifting polarity balance, confirming α1-ARs mediate slow dynamics. Prazosin did not affect fast transients (Supplemental Fig. 5). Propranolol (β-AR antagonist) affected neither fast nor slow activities (Supplemental Fig. 5), demonstrating β-ARs do not contribute.

Pathway-Specific Effects on Purkinje Cell Synaptic Inputs

Ex vivo patch-clamp recordings provided pathway-specific validation. For climbing fiber (CF) stimulation (Fig. 2q-r), yohimbine increased CF-evoked excitatory postsynaptic currents (EPSCs), while expressing plasma membrane Ca^{2+} ATPase (PMCA)[41,42] in BG via adeno-associated virus (AAV) injection—which enhances Ca^{2+} export from astrocytes and reduces astrocytic Ca^{2+} activity—significantly reduced responses.

For parallel fiber (PF) stimulation (Fig. 2s-t), prazosin decreased PF-evoked EPSCs, and PMCA expression further reduced them.

Immunohistochemistry confirmed α2-ARs were highly expressed in BG, α1-ARs co-labeled with S100β in astrocytes, and α1-ARs were adjacent to PF terminals marked by VGLUT1 (Supplemental Fig. 6). Conditional α1-AR knockout in BG nearly eliminated α1-AR-evoked Ca^{2+} activity in cerebellar cortex[43], demonstrating functional localization to astrocytes adjacent to PF synapses. Prazosin did not affect CF-evoked EPSCs, consistent with α1-AR localization to PF pathways[44,45], while yohimbine significantly reduced PF-evoked EPSCs during PF stimulation (Supplemental Fig. 5). Given that CF signals are classically viewed as teaching signals[46-48] because they trigger complex spikes that suppress simple spikes from PF input, the observation that yohimbine enhances CF input while simultaneously suppressing PF input suggests α2-AR signaling coordinates the balance between these pathways.

The temporal properties align with BG dynamics: granule cells providing PF input generate sparse, sustained contextual signals[49] matching slow BG activities, while CF inputs deliver sharp,

temporally precise instructive signals[49] matching fast transients. This suggests slow α1-AR-mediated signals regulate contextual PF pathways, while fast α2-AR-mediated signals regulate instructive CF pathways.

Fast and Slow BG Activities Differentially Regulate Reward-Guided Behavior

We selectively perturbed fast α2-AR-dependent or slow α1-AR-dependent dynamics during the reward-guided task. For fast activity disruption, we expressed ChETA-EYFP[50,51] in BG and delivered optogenetic stimulation upon sensor breaks in trigger or reward zones during Days 1-10, followed by no-stimulation recovery (Days 11-15; Fig. 3a). This produced specific impairments: ChETA animals showed more trigger zone crossings as more sensor breaks, less reward licking, lower rewarded-to-total lick ratio, and increased trigger-to-reward time, consistent with impaired reinforcement learning and altered reward behavior. ChETA effects depended on task alignment: stimulation delivered every 10 s independently of behavior produced no performance changes (Supplemental Fig. 7), demonstrating fast BG activity influence requires coupling to task-relevant cues.

For slow signal disruption, mice received bilateral prazosin or vehicle before daily sessions from Day 1-10, followed by no-drug recovery (Day 11-15; Fig. 3b). Prazosin impaired distinct behaviors: longer trigger-to-reward time, longer reward-to-trigger time, prolonged trial duration, and reduced trial number—all related to switching between reward-seeking and rewarding states.

BG Ca^{2+} dynamics thus have dissociable roles: fast α2-AR-modulated activities regulate temporally precise cue responses, reinforcement learning, and reward behavior, while slow α1-AR-mediated activities support sustained task structure and behavioral transitions.

Actor-Critic Network Recapitulates Multitimescale Astrocytic Computation

Our results reveal division of labor where astrocytes modulate synaptic transmission at CF and PF inputs to Purkinje cells, differentially regulating instructive versus contextual pathways through distinct temporal modes. Disrupting astrocytic activities impairs learning and performance (Fig. 3), demonstrating astrocytes actively adjust circuit dynamics to guide behavior.

This modulatory architecture resembles actor-critic frameworks in reinforcement learning[52-55]: actors generate actions while critics evaluate states to modulate learning, improving performance without directly controlling behavior[53]. We hypothesized this represents a general organizational principle: neurons act as actors generating outputs, while astrocytes act as critics evaluating states and modulating computation.

If this captures neuron-astrocyte computation, then artificial networks implementing explicit actor-critic separation should spontaneously develop the multitimescale, heterogeneous astrocytic dynamics we observe biologically.

We trained a biologically-inspired actor-critic network on the spatial sequence task using Proximal Policy Optimization[56,57] (Fig. 4a). The network comprised 16 neuronal units (actor) generating actions and 16 astrocytic units (critic) estimating state values. Neuron-astrocyte interactions

followed tripartite synapse principles[52,58], where astrocytes bidirectionally modulate synaptic transmission—capturing key properties of biological astrocytic regulation.

Trained neuronal and astrocytic networks occupied distinct state space regions (Fig. 4b), demonstrating functional differentiation. Critically, multitimescale astrocytic signals emerged spontaneously through training (Fig. 4c). Pre-training activity was minimal, matching Day 1 recordings. Post-training revealed fast transients at trigger zone entry (sensory cue) and slow ramping with polarity reversals during state transitions—closely resembling biological state-change dynamics (Fig. 1f).

Individual astrocytes developed heterogeneous response profiles (Fig. 4d), including peak-like and trough-like responses—recapitulating diversity observed in Bergmann glia (Fig. 1e,g). The model reproduced the integrative property: individual astrocytes exhibiting peak-like and trough-like signals that, when summed across the population, generated state-change dynamics—providing evidence that state transitions emerge from distributed astrocytic integration, as recording data suggests (Supplemental Fig. 3).

To assess whether astrocyte-inspired architectures confer computational advantages, we compared our astrocyte-neuron network (AsNeN) with traditional architectures including vanilla Recurrent Neural Networks (vRNN) and Long Short-Term Memory networks (LSTM). The AsNeN network outperformed both traditional architectures across multiple performance metrics (Fig. 4e-f). AsNeN achieved consistently higher moving-average reward and success rates compared to vRNN and LSTM, converged faster in both metrics, and exhibited greater post-convergence stability with minimal fluctuations compared to the oscillatory behavior of traditional networks.

Actor-critic architectures thus spontaneously develop multitimescale, heterogeneous astrocytic computation observed biologically. Convergence between biological observations and computational theory, combined with superior performance over traditional architectures, suggests actor-critic computation may be a general organizational strategy for neuron-astrocyte networks during behavior.

Discussion

Astrocytes as Multilevel Computational Elements

Our findings establish that astrocytes perform computation by transforming univariate neuromodulatory inputs into multivariate, pathway-specific circuit regulation across distinct timescales. Three features define this computation: temporal decomposition (NE parsed into slow and fast activities through α1-ARs and α2-ARs), spatial specificity (each activity selectively targets distinct synaptic pathways—slow activities modulate PF inputs, fast activities regulate CF inputs), and emergent integration (population-level summation of opposing peak and trough signals generates state-change dynamics). This integrative mechanism receives further support from our actor-critic network simulations, which spontaneously reproduced peak and trough signals that sum to generate state-change dynamics, strengthening the hypothesis that state transitions arise from distributed astrocytic computation[57,58].

Complementary Computational Properties

Neurons and astrocytes differ fundamentally in their computational properties. Neurons generate action potentials, providing high-bandwidth, temporally precise, point-to-point signaling[1-3], while astrocytes lack voltage-gated sodium channels[59,60] and rely on graded Ca^{2+} elevations[4,5] that provide lower-bandwidth, temporally extended, spatially integrative signaling. Neurons maintain high membrane resistance for precise signal propagation, while astrocytes exhibit low resistance and extensive gap junction coupling[6-8], creating networks that integrate information over large territorial domains and influence tens of thousands of synapses[6-8]. This suggests dual computational architecture: neurons excel at temporally precise, spatially specific computation through electrical signaling; astrocytes excel at temporally extended, spatially integrative computation through chemical signaling[61,62].

While we focus on norepinephrine, astrocytes express diverse receptors for dopamine, serotonin, and acetylcholine[63-65], many featuring receptor subtypes with distinct kinetics (e.g., fast nicotinic versus slow muscarinic receptors[66]). If astrocytes simultaneously parse multiple neuromodulators into different timescale activities, they could generate a combinatorially large control space from few inputs. The fast Ca^{2+} transients we observe may not be unique to Bergmann glia—recent studies identified fast astrocytic signals in cortex[16-18], hippocampus[67,68], and striatum[69]. Our key contribution is demonstrating that fast signals emerge with learning, show behavioral tuning, and causally regulate behavior. The cerebellar cortex may provide favorable conditions due to climbing fibers delivering precisely timed instructive signals[22,23], but regions with temporally structured inputs may reveal fast signals under appropriate behavioral conditions. Disrupting each timescale produces distinct behavioral deficits—fast signal loss impairs event-triggered responses while slow signal loss impairs state transitions—demonstrating both activities are necessary and non-redundant.

Actor-Critic Computation: A Unifying Framework

Our computational modeling reveals that multitimescale, heterogeneous astrocytic dynamics emerge naturally from actor-critic architectures where neurons act as actors generating behavioral outputs and astrocytes act as critics evaluating states to modulate learning. Implementing neuron-astrocyte interactions following tripartite synapse principles[50,56] was sufficient to recapitulate biological astrocytic computation, suggesting that bidirectional astrocyte-synapse communication combined with reinforcement learning naturally produces the temporal decomposition and population integration we observe experimentally. The spontaneous emergence of biologically realistic dynamics in artificial networks trained on the same task provides strong evidence that astrocytes perform computation rather than merely reacting to neuronal activity—if astrocytes were passive responders, their dynamics would mirror neuronal patterns rather than developing distinct temporal profiles optimized for state evaluation.

Implications for Artificial Intelligence

Actor-critic architectures incorporating astrocyte-inspired critic networks can solve complex tasks while spontaneously developing multitimescale dynamics observed in biological systems. Our direct comparison demonstrates that the astrocyte-neuron network outperforms traditional

architectures including vanilla Recurrent Neural Networks and Long Short-Term Memory networks in learning efficiency, performance, and stability. This suggests design principles for artificial neural networks[70-73]: temporal decomposition layers that automatically parse inputs into fast (event-level) and slow (state-level) components; low-precision, high-integration units with large receptive fields for detecting population patterns[71]; and neuromodulatory decoding modules that transform global signals into pathway-specific regulation. Such architectures could address continual learning challenges by balancing stability (slow-timescale units maintaining task structure) and plasticity (fast-timescale units enabling rapid adaptation), improve credit assignment through population-level error signals, and enhance meta-learning through multidimensional control spaces. Recognizing astrocytes as multilevel computational elements may fundamentally reshape understanding of information processing in biological and artificial systems, potentially leading to more capable artificial systems leveraging complementary computational architectures.

Methods

Animals

All procedures were approved by the University of Alabama at Birmingham Institutional Animal Care and Use Committee and followed NIH guidelines. Male and female mice (8-12 weeks old) were group-housed on a 12h light/dark cycle with ad libitum access to food and water except during behavioral testing. Experimental groups included: Aldh1l1-CreERT2 mice (Jackson Laboratory) crossed with Rosa26-CAG-GCaMP6f reporter mice for astrocyte-specific calcium imaging; wild-type C57BL/6J mice for GRAB_NE sensor experiments and pharmacological manipulations; and Aldh1l1-CreERT2 mice for optogenetic manipulations. Tamoxifen (75 mg/kg body weight, i.p.) was administered three times with 2-day intervals between injections to induce Cre recombination in Aldh1l1-CreERT2 lines. Animals were randomly assigned to experimental groups, and experimenters were blinded to group assignment during behavioral testing and analysis.

Viral Constructs and Stereotactic Surgery

For GCaMP6f expression, Aldh1l1-CreERT2; Rosa26-CAG-GCaMP6f mice were used without additional viral injection. For GRAB_NE expression, pAAV-hSyn-GRAB_NE1m (Addgene) was injected into cerebellar cortex lobules III and left and right simplex. For optogenetic experiments, pAAV-Ef1a-DIO-ChETA-EYFP or pAAV-Ef1a-DIO-EYFP control (Addgene) was injected into left and right simplex regions. For PMCA overexpression, pZac2.1-GfaABCD-nCherry-hPMCA2w/b (AAV5, Addgene) was injected into cerebellar cortex lobules III and left and right simplex.

Mice were anesthetized with isoflurane (1.5-2% in O_2) and placed in a stereotactic frame. Body temperature was maintained at 37°C using a feedback-controlled heating pad. The skull was exposed, and small craniotomies were made above target coordinates. All coordinates were referenced to Lambda. For cerebellar lobule III: AP −2.4 mm from Lambda, ML 0 mm, DV 1.0 mm from brain surface. For cerebellar simplex: AP −2.4 mm from Lambda, ML ±2.3 mm, DV 1.0 mm from brain surface. Virus (200 nl per site) was injected at 50 nl/min using a micropump, and

the needle remained in place for 10 min post-injection before slow retraction. For fiber photometry, fiber optic cannulas (200 µm core diameter, 0.37 NA, RWD Life Science) were implanted above injection sites and secured with dental cement. For pharmacological experiments, guide cannulas (RWD Life Science) were implanted bilaterally at target sites. For combined photometry and drug delivery, Multiple Fluid Injections Cannulas (200 µm, 0.37 NA, Doric Lenses) were used. Animals recovered for 3-4 weeks before behavioral training to allow optimal expression.

Behavioral Task

Mice were water-restricted to 1 ml per day starting 2 days before behavioral experiments and trained on a custom spatial sequence task in a rectangular arena (60 × 8 cm) with distinct trigger and reward zones (10 × 10 cm each) at opposite ends. Infrared (IR) beam-break sensors detected zone entries/exits. The task required: (1) entering the trigger zone, which activated an auditory tone (8 kHz, 0.2 s duration), (2) exiting the trigger zone and traversing to the reward zone, (3) entering the reward zone and waiting 2 s, after which a liquid sucrose reward (5% w/v, 4 µl) was delivered via a solenoid-controlled lick port, (4) consuming reward during a 4 s window, then (5) returning to the trigger zone to initiate the next trial. Lick events were detected by the breaking of the IR sensor at the port. Training consisted of 30 min daily sessions for 15 consecutive days. Day 1 served as a no-reward baseline to assess exploratory behavior. Behavioral metrics included: number of successful trials, trigger-to-reward zone time, reward-to-trigger zone time, trial duration, licking bouts with/without reward, and rewarded/total lick ratio.

Fiber Photometry

Fiber photometry was performed using an RWD R821/FR-21 Tricolor Multichannel Fiber Photometry System. Blue LED light (470 nm) and an isosbestic wavelength control (410 nm) were both modulated at 15 Hz during recording and delivered through the implanted fiber. Signals were digitized at 1 kHz and synchronized with behavioral events via TTL pulses from IR sensors.

Data were analyzed in MATLAB and Python. For GCaMP6f signals, raw fluorescence (F) was calculated as $\Delta F/F = (F_{470} - F_{410})/F_{410}$ after correcting for photobleaching using exponential fitting. For GRAB_NE signals, the 0-200 s period of each session was used as baseline for normalization, and then a −2 to +2 s window around each event was used for event-specific baseline correction. Signals were then normalized to z-scores for cross-session comparisons. Trained data was defined as the day the animal completed the most trials after day 7.

Slow signal classification was performed based on activity patterns during licking onset. Recording sites were classified as trough, state-change, or peak signals based on the presence of troughs or peaks within a −1 to +1 s window around licking onset. Sites showing clear troughs were classified as trough signals, sites showing clear peaks were classified as peak signals, and sites showing neither prominent peaks nor troughs were classified as state-change signals.

Pharmacology

For receptor antagonism during photometry recordings, drugs were infused through the Multiple Fluid Injections Cannulas (Doric Lenses) used for combined optical recording and drug delivery

after 10 days of training. Yohimbine hydrochloride (α2-AR antagonist, Tocris), prazosin hydrochloride (α1-AR antagonist, Tocris), propranolol hydrochloride (β-AR antagonist, Tocris), or vehicle (aCSF) were prepared at 1 mM concentration and infused at 1 µl per recording site at 250 nl/min. After infusion completion, the injector remained in place for 5 min to prevent backflow. Behavioral sessions and recordings began 10 min post-infusion. Each animal received all treatments in counterbalanced order with ≥3 days washout between sessions.

For behavioral manipulation, prazosin (1 mM in aCSF, 1 µl per side) or vehicle was infused bilaterally 10 min before daily training sessions (Days 1-10) via implanted guide cannulas using the same infusion protocol (250 nl/min, 5 min hold time). Infusions ceased on Days 11-15 to assess recovery.

Optogenetic Manipulation

For fast signal disruption, Aldh1l1-CreERT2 mice expressing ChETA-EYFP or control EYFP in Bergmann glia received bilateral fiber optic implants above cerebellar simplex. Laser stimulation (473 nm, 10 mW, 50 ms pulses) was triggered automatically by IR beam breaks when animals entered or exited the trigger zone or reward zone during Days 1-10 of training. To prevent repeated stimulation during prolonged zone occupancy, a 5 s refractory period was implemented following each stimulation event. Stimulation ceased on Days 11-15 to assess recovery. Control experiments used the same stimulation protocol but delivered independently of behavior (every 10 s) or used EYFP-expressing control animals with behavior-triggered stimulation.

Ex Vivo Electrophysiology

Acute cerebellar slices (300 µm) were prepared from wild-type mice (3-8 weeks old) or mice expressing PMCA (4 weeks post-AAV injection). Mice were deeply anesthetized with a ketamine and xylazine mixture, and transcardially perfused with ice-cold cutting solution containing (in mM): 87 NaCl, 2.5 KCl, 0.5 $CaCl_2$, 7 $MgCl_2$, 1.25 NaH_2PO_4, 25 $NaHCO_3$, 25 glucose, and 75 sucrose, bubbled with 95% O_2/5% CO_2. The brain was rapidly removed and cut transversely using a vibratome (VT1200S, Leica Microsystems). Slices were transferred to oxygenated aCSF at 32°C for 30 min, then allowed to recover for 1 h at room temperature before recordings. Recording aCSF contained (in mM): 119 NaCl, 2.5 KCl, 2.5 $CaCl_2$, 1.3 $MgCl_2$, 1.3 NaH_2PO_4, 26 $NaHCO_3$, and 20 glucose.

Whole-cell voltage-clamp recordings were obtained from visually identified Purkinje cells using an upright microscope (Axio Examiner.D1, Zeiss) with IR-DIC optics. Individual slices were transferred to a submerged chamber and continuously perfused with normal oxygenated aCSF at room temperature. Patch pipettes (3-4 MΩ) were filled with internal solution containing (in mM): 120 Cs-gluconate, 17.5 CsCl, 10 Na-HEPES, 4 Mg-ATP, 0.4 Na-GTP, 10 Na_2-creatine phosphate, 0.2 Na-EGTA (290-300 mOsm, pH 7.3). Cells were held at −60 mV. Cells with series resistances above 25 MΩ were discarded, and cells were also excluded if any whole-cell parameter (i.e., Cm, Ri, Rs) changed by ≥20% during recordings.

For climbing fiber and parallel fiber stimulation, aCSF-filled pipettes connected to an isolated stimulator (ISO-Flex, AMPI) were used to deliver electrical stimulation every 20 s. For climbing

fiber stimulation, the stimulating electrode was placed in the granule cell layer to elicit all-or-none climbing fiber EPSCs. For parallel fiber stimulation, the electrode was placed in the molecular layer to evoke graded EPSCs. Stimulus intensity was adjusted to ensure EPSCs were clamped. Synaptic responses were recorded during baseline (0-300 s), drug application (300-1100 s), and washout (1100-1880 s). Prazosin (10 μM) or yohimbine (15 μM) was bath applied during the drug application period.

Data were acquired using TI Workbench software with a MultiClamp 700B amplifier (Molecular Devices), filtered at 2 kHz, and digitized at 10 kHz with ITC-18 A/D-D/A interfaces (Instrutech). Analysis was performed using TI Workbench and Python.

Immunohistochemistry

Mice were deeply anesthetized and transcardially perfused with 4% paraformaldehyde in PBS. Brains were removed and post-fixed overnight in 4% PFA at 4°C. Coronal brain sections were cut at 60 μm using a vibratome. Sections were permeabilized with 0.25% Triton X-100 for 2 h at room temperature and blocked with 10% normal goat serum for 1 h. Sections were incubated at 4°C overnight with blocking solution containing the following primary antibodies: rabbit anti-α1-AR (1:500, Invitrogen), rabbit anti-α2-AR (1:500, Proteintech), mouse anti-S100β (1:500, Invitrogen), guinea pig anti-VGLUT1 (1:500, MilliporeSigma). After primary antibody incubation, sections were rinsed with PBS three times for 10 min each and incubated for 2 h at room temperature with Alexa Fluor-conjugated secondary antibodies (1:500, Jackson ImmunoResearch Laboratories). Sections were coverslipped with Vectashield mounting medium (Vector Laboratories). Confocal images were acquired on a Zeiss LSM 880 microscope using 20x (overview) and 40x oil-immersion objectives (cellular detail). Colocalization analysis was performed using ImageJ with JACoP plugin.

Bio-Inspired Neuron-Astrocyte Network Simulation

To simulate and investigate how astrocyte activity supports and modulates behavior, we designed an artificial task in which an agent equipped with a biologically inspired neuron–astrocyte network is trained with reinforcement learning (RL) to solve the task in a manner analogous to the experiment.

Task environment: spatial sequence reward task

We built an artificial spatial sequence reward task that captures the key structure of the mouse experiment. In the task, the agent needs to move from a trigger zone to a reward zone and obtain reward after a delay. Each trial is implemented as a discrete-time Markov decision process with four states: a trigger zone, a transit state, a reward zone, and the terminal success or failure. At each time step, the agent can select one of two actions, move forward or wait/stay.

At the beginning of a trial, the agent typically starts in the trigger zone, where a brief cue is initiated, representing the sensory cue in the experiment. If the agent chooses the move action, the state transitions into the transit phase that lasts a fixed number of time steps. The agent needs to keep move in this phase and then can enter the reward zone. In the reward zone, the agent must emit the wait action for enough consecutive time steps to receive the reward. This sustained waiting period mimics licking at the reward spout after a fixed delay. When the accumulated hold time exceeds the reward delay, the agent receives a unit reward (+1), followed by a brief post-reward

period after which the episode terminates in a success state. If the agent fails to obtain reward before a hard maximum trial length, the episode terminates in failure. During the task, each time step carries a small step cost (−0.01), so longer or inefficient trajectories are mildly penalized, encouraging the agent to move efficiently and to wait only as long as necessary.

At each time step, the agent receives a three-dimensional observation: the first component is a constant bias, i.e., +1; the second is the cue channel, which is transiently active after entering the trigger zone, i.e., +1 or 0; and the last component indicates whether reward has been delivered, i.e., +1 or 0. This minimal environment preserves the temporal structure of the experiment, namely cue-gated departure from the trigger zone, delayed reward contingent on sustained waiting, and time-limited trials. Moreover, it remains simple enough and facilitate a systematical analysis the learned neuron–astrocyte dynamics.

Bio-inspired neuron-astrocyte network

The agent was implemented with a biologically inspired recurrent neuron-astrocyte network with standard linear input and output layers. The recurrent network is an extension of the previous neuron-astrocyte network model from (Gong et al., 2024), which is built based on the tripartite synapse structure of neuron and astrocyte interactions. The network consists of neuronal activities $x_t \in \mathbb{R}^{N_x}$, astrocytic activities $z_t \in \mathbb{R}^{N_z}$, and a flattened synaptic weight vector w_t, which is reshaped into a recurrent weight matrix W_t. In the simulations reported here, we set $N_x = N_z = 16$, and the synaptic weight w_t is sparsely initialized.

Let o_t denote the 3-dimensional observation from the task environment at time t. The updates for neurons, synapses, and astrocytes are given by the following system:

$$x_t = (1-\gamma)x_{t-1} + \gamma(\sigma(W_{t-1}x_{t-1}) + \sigma(W_{in,x}o_{t-1})$$

$$w_t = (1-\gamma)w_{t-1} + \gamma\left(\sigma\left((x_{t-1}x_{t-1}^T) \odot C\right) + tanh(Dz_{t-1})\right)$$

$$z_t = (1-\gamma)z_{t-1} + \gamma\tau\left(tanh(Fz_{t-1}) + H \odot tanh(x_{t-1}x_{t-1}^T) + tanh(W_{in,z}o_{t-1}\right)$$

Here, σ is the sigmoid activation and tanh is the hyperbolic tangent activation function. γ is the fixed discretization step, and $\tau < 1$ denotes the slow astrocyte dynamics relative to neurons. The matrix C parameterizes element-wise gains on the outer product $x_t\, x_t^T$, D maps astrocyte activity to synaptic modulation, H maps synaptic activity back to astrocytes, and F captures astrocyte–astrocyte coupling. The matrices $W_{in,x}$ and $W_{in,z}$ inject task observations into neurons and astrocytes, respectively. This architecture implements the biologically tripartite feedback loop in which neuronal activity drives synaptic and astrocytic updates, and astrocytes feed back to synapses and neurons to provide slow, context-dependent modulation (see more details of the network model in Gong et al., 2024)[56].

Reinforcement learning agent and training procedure

We implemented this neuron-astrocyte-network-embedded agent in the reinforcement learning framework, i.e., the proximal policy optimization (PPO), to solve the defined spatial sequence reward task. We define separate readouts to generate the policy and value functions. For the policy, we interpret the neuronal layer as the action-selecting module, but allow astrocytes to modulate it. Specifically, the action probabilities over the two discrete actions (move, wait) are given by a SoftMax readout $\pi(a_t \mid o_t) = \mathrm{SoftMax}(W_\pi\ x_t)$. The value function is outputted directly from astrocyte activity via a linear map $V(o_t) = W_v\ z_t$, reflecting the role of astrocytes as a slower critic-like integrator of reward history and task context.

All the parameters in the neuron-astrocyte network, policy and value readouts are trained with Adam using multiple PPO epochs per update and minibatches. During the training, we track the

obtained reward and the binary success flag for each episode. To assess the convergence and performance, we compute the average reward and success rate over a moving window of multiple episodes. All simulations are implemented in Python using PyTorch with fixed random seeds and deterministic settings to enhance reproducibility. The setting and hyperparameters in the simulation are summarized in Table 1.

Table 1. Simulation settings

Category	Parameter	Value
Network	Neuron units (N_x)	16
	Astrocyte units (N_z)	16
	Discretization step (γ)	0.01
PPO / RL	Policy learning rate	0.001
	PPO clip	0.2
	Entropy coefficient	0.05
	Rollout horizon (T)	64
	PPO epochs per update	16
	Minibatch size	64
	Training episodes	10,000
Environment	Reward delay	2-8
	Max trial length	20 steps

Analysis of trained network dynamics

After training, we analyze both behavioral performance of the agent and internal dynamics of the neuron–astrocyte network, focusing on astrocyte activity and its interaction with neuronal activity. We generate rollouts using the trained policy in a fixed environment configuration. For each time step, we record the full internal state [x_t, w_t, z_t], the environment state label, the chosen action, and key task events (i.e., enter trigger zone, leave trigger zone, enter reward zone, receive reward).

We separate the recorded states into neuronal activity x_t and astrocyte activity z_t and examine their dynamics on single and multiple successful trials. To obtain low-dimensional summaries, we perform standard principal component analysis (PCA) on x_t and z_t. For single trials, PCA is computed on that trial's activity; for multi-trial analyses, we concatenate activity across several successful trials and compute a shared PCA basis, then project each trial into this basis. We plot the first principal component as a function of time, aligned to behavioral events and overlaid with the action sequence. To visualize the full astrocyte population, we also plot raw states of individual astrocytes across time and units. These visualizations reveal both fast, event-locked components and slower ramps or plateaus aligned to reward-related epochs, which is aligned with the experimental findings.

As a control, we also perform the same analyses on untrained (randomly initialized) networks underlying the same environment. We compare trained and untrained dynamics, which demonstrates how neuronal and astrocytic activities are reshaped in the reinforcement learning process to solve the task.

Learning performance comparison with standard network architectures

We further evaluate the bio-inspired neuron-astrocyte network by comparing it with two standard recurrent architectures, a vanilla RNN and an LSTM, applied to the same artificial task. Both baseline networks use a hidden dimension of 128, having a parameter count comparable to that of the neuron-astrocyte network. All networks are trained using the same PPO framework with the same hyperparameter settings. Then, we compare the learning performance by monitoring two metrics: the episodic reward and the success rate, each reported as a moving average over a 100-episode window. The results show that the neuron-astrocyte network outperforms these two networks by having relatively higher reward and success rate throughout training. Moreover, its learning curves converge faster and exhibit much less variance, which indicates better learning stability compared to the standard recurrent baselines.

Statistical Analysis

All statistical analyses were performed in MATLAB, Python, and GraphPad Prism. Data are presented as mean ± SEM unless otherwise noted. Sample sizes were determined based on pilot experiments and previously published studies. Normal distribution was assessed using Shapiro-Wilk tests. For normally distributed data, parametric tests were used: paired or unpaired t-tests for two-group comparisons, one-way or two-way ANOVA with Tukey's or Sidak's post-hoc tests for multiple comparisons. For non-normally distributed data or small sample sizes, non-parametric tests were used: Mann-Whitney U test, Wilcoxon signed-rank test, or Kruskal-Wallis test with Dunn's post-hoc. Repeated measures ANOVA was used for longitudinal behavioral data across training days. Correlations were assessed using Spearman's rank correlation coefficient for non-parametric data or when examining monotonic relationships. Statistical significance was set at $\alpha = 0.05$. No statistical methods were used to predetermine sample sizes, but sample sizes are consistent with standards in the field.

Figure Preparation

Schematic illustrations and diagrams were created using BioRender (biorender.com). Three-dimensional brain illustrations and anatomical visualizations were generated using HERBS (Histological E-data Registration in rodent Brain Spaces) in conjunction with the Allen Mouse Brain Common Coordinate Framework[74,75]. Confocal microscopy images were processed using ImageJ/Fiji with brightness and contrast adjustments applied uniformly across all comparison groups.

Data and Code Availability

All data supporting the findings of this study and custom analysis code are available from the corresponding author upon reasonable request. Computational models will be deposited in a public repository (GitHub) upon publication.

References

1. Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol. 1952;117(4):500-544.
2. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys. 1943;5:115-133.
3. Carandini M, Heeger DJ. Normalization as a canonical neural computation. Nat Rev Neurosci. 2012;13(1):51-62.
4. Bazargani N, Attwell D. Astrocyte calcium signaling: the third wave. Nat Neurosci. 2016;19(2):182-189.
5. Volterra A, Liaudet N, Savtchouk I. Astrocyte Ca^{2+} signalling: an unexpected complexity. Nat Rev Neurosci. 2014;15(5):327-335.
6. Bushong EA, Martone ME, Jones YZ, Ellisman MH. Protoplasmic astrocytes in CA1 stratum radiatum occupy separate anatomical domains. J Neurosci. 2002;22(1):183-192.
7. Halassa MM, Fellin T, Haydon PG. The tripartite synapse: roles for gliotransmission in health and disease. Trends Mol Med. 2007;13(2):54-63.
8. Araque A, Parpura V, Sanzgiri RP, Haydon PG. Tripartite synapses: glia, the unacknowledged partner. Trends Neurosci. 1999;22(5):208-215.
9. Paukert M, Agarwal A, Cha J, et al. Norepinephrine controls astroglial responsiveness to local circuit activity. Neuron. 2014;82(6):1263-1270.
10. Savtchouk I, Volterra A. Gliotransmission: Beyond black-and-white. J Neurosci. 2018;38(1):14-25.
11. Cornell-Bell AH, Finkbeiner SM, Cooper MS, Smith SJ. Glutamate induces calcium waves in cultured astrocytes: long-range glial signaling. Science. 1990;247(4941):470-473.
12. Khakh BS, McCarthy KD. Astrocyte calcium signaling: from observations to functions and the challenges therein. Cold Spring Harb Perspect Biol. 2015;7(4):a020404.
13. Shigetomi E, Bushong EA, Haustein MD, et al. Imaging calcium microdomains within entire astrocyte territories and endfeet with GCaMPs expressed using adeno-associated viruses. J Gen Physiol. 2013;141(5):633-647.
14. Araque A, Carmignoto G, Haydon PG, et al. Gliotransmitters travel in time and space. Neuron. 2014;81(4):728-739.
15. Poskanzer KE, Yuste R. Astrocytes regulate cortical state switching in vivo. Proc Natl Acad Sci USA. 2016;113(19):E2675-E2684.
16. Bindocci E, Savtchouk I, Liaudet N, et al. Three-dimensional Ca^{2+} imaging advances understanding of astrocyte biology. Science. 2017;356(6339):eaai8185.
17. Stobart JL, Ferrari KD, Barrett MJP, et al. Cortical circuit activity evokes rapid astrocyte calcium signals on a similar timescale to neurons. Neuron. 2018;98(4):726-735.
18. Semyanov A, Verkhratsky A. Astrocytic processes: from tripartite synapses to the active milieu. Trends Neurosci. 2021;44(10):781-792.
19. Nimmerjahn A, Mukamel EA, Schnitzer MJ. Motor behavior activates Bergmann glial networks. Neuron. 2009;62(3):400-412.
20. Kanemaru K, Sekiya H, Xu M, et al. In vivo visualization of subtle, transient, and local activity of astrocytes using an ultrasensitive Ca^{2+} indicator. Cell Rep. 2014;8(1):311-318.
21. Agarwal A, Wu PH, Hughes EG, et al. Transient opening of the mitochondrial permeability transition pore induces microdomain calcium transients in astrocyte processes. Neuron. 2017;93(3):587-605.

22. Ito M. Cerebellar circuitry as a neuronal machine. Prog Neurobiol. 2006;78(3-5):272-303.
23. Albus JS. A theory of cerebellar function. Math Biosci. 1971;10(1-2):25-61.
24. Eccles JC, Llinás R, Sasaki K. The excitatory synaptic action of climbing fibres on the Purkinje cells of the cerebellum. J Physiol. 1966;182(2):268-296.
25. Marr D. A theory of cerebellar cortex. J Physiol. 1969;202(2):437-470.
26. Palay SL, Chan-Palay V. Cerebellar Cortex: Cytology and Organization. Springer-Verlag; 1974.
27. De Zeeuw CI, Hoebeek FE, Bosman LW, et al. Spatiotemporal firing patterns in the cerebellum. Nat Rev Neurosci. 2011;12(6):327-344.
28. Balakrishnan, S., & Bellamy, T. C. Depression of parallel and climbing fiber transmission to Bergmann glia is input specific and correlates with increased precision of synaptic transmission. Glia. 2009;*57*(4):393-401.
29. Beierlein M, Regehr WG. Brief bursts of parallel fiber activity trigger calcium signals in Bergmann glia. J Neurosci. 2006;26(26):6958-6967.
30. Cui G, Jun SB, Jin X, et al. Concurrent activation of striatal direct and indirect pathways during action initiation. Nature. 2013;494(7436):238-242.
31. Gunaydin LA, Grosenick L, Finkelstein JC, et al. Natural neural projection dynamics underlying social behavior. Cell. 2014;157(7):1535-1551.
32. Taniguchi H, He M, Wu P, et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron. 2011;71(6):995-1013.
33. Chen TW, Wardill TJ, Sun Y, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature. 2013;499(7458):295-300.
34. Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci. 2005;28:403-450.
35. Chandler DJ, Gao WJ, Waterhouse BD. Heterogeneous organization of the locus coeruleus projections to prefrontal and motor cortices. Proc Natl Acad Sci USA. 2014;111(18):6816-6821.
36. Schwarz LA, Miyamichi K, Gao XJ, et al. Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. Nature. 2015;524(7563):88-92.
37. Carey MR, Regehr WG. Noradrenergic control of associative synaptic plasticity by selective modulation of instructive signals. Neuron. 2009;62(1):112-122.
38. Jirenhed DA, Bengtsson F, Hesslow G. Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. J Neurosci. 2007;27(10):2493-2502.
39. Schonewille M, Gao Z, Boele HJ, et al. Reevaluating the role of LTD in cerebellar motor learning. Neuron. 2011;70(1):43-50.
40. Feng J, Zhang C, Lischinsky JE, et al. A genetically encoded fluorescent sensor for rapid and specific in vivo detection of norepinephrine. Neuron. 2019;102(4):745-761.
41. Yu X, Taylor AMW, Nagai J, et al. Reducing astrocyte calcium signaling in vivo alters striatal microcircuits and causes repetitive behavior. Neuron. 2018;99(6):1170-1187.
42. Strehler EE. Plasma membrane calcium ATPases: from generic Ca2+ sump pumps to versatile systems for fine-tuning cellular Ca2+. Biochem Biophys Res Commun. 2015;460(1):26-33.
43. Ye L, Orynbayev M, Zhu X, et al. Ethanol abolishes vigilance-dependent astroglia network activation in mice by inhibiting norepinephrine release. Nat Commun. 2020;11(1):6157.
44. Eccles JC, Llinás R, Sasaki K. The excitatory synaptic action of climbing fibres on the Purkinje cells of the cerebellum. J Physiol. 1966;182(2):268-296.
45. Albus JS. A theory of cerebellar function. Math Biosci. 1971;10(1-2):25-61.

46. Ito M. Cerebellar long-term depression: characterization, signal transduction, and functional roles. Physiol Rev. 2001;81(3):1143-1195.
47. Garcia-Garcia MG, Kapoor A, Akinwale O, et al. A cerebellar granule cell-climbing fiber computation to learn to track long time intervals. Neuron. 2024;112(16):2215-2230.
48. Gunaydin LA, Yizhar O, Berndt A, et al. Ultrafast optogenetic control. Nat Neurosci. 2010;13(3):387-392.
49. Lin JY, Lin MZ, Steinbach P, Tsien RY. Characterization of engineered channelrhodopsin variants with improved properties and kinetics. Biophys J. 2009;96(5):1803-1814.
50. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. MIT Press; 2018.
51. Mnih V, Badia AP, Mirza M, et al. Asynchronous methods for deep reinforcement learning. Proc Int Conf Mach Learn. 2016;48:1928-1937.
52. Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. Nature. 2016;529(7587):484-489.
53. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997;275(5306):1593-1599.
54. Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms. arXiv:1707.06347. 2017.
55. Schulman J, Levine S, Abbeel P, et al. Trust region policy optimization. Proc Int Conf Mach Learn. 2015;37:1889-1897.
56. Gong L, Pasqualetti F, Papouin T, Ching SN. Astrocytes as a mechanism for contextually-guided network dynamics and function. PLoS Comput Biol. 2024;20(5):e1012186.
57. Oberheim NA, Takano T, Han X, et al. Uniquely hominid features of adult human astrocytes. J Neurosci. 2009;29(10):3276-3287.
58. Nedergaard M, Ransom B, Goldman SA. New roles for astrocytes: redefining the functional architecture of the brain. Trends Neurosci. 2003;26(10):523-530.
59. Verkhratsky A, Nedergaard M. Physiology of astroglia. Physiol Rev. 2018;98(1):239-389.
60. Khakh BS, Sofroniew MV. Diversity of astrocyte functions and phenotypes in neural circuits. Nat Neurosci. 2015;18(7):942-952.
61. Savtchouk I, Volterra A. Gliotransmission: Beyond black-and-white. J Neurosci. 2018;38(1):14-25.
62. Allen NJ, Lyons DA. Glia as architects of central nervous system formation and function. Science. 2018;362(6411):181-185.
63. Bekar LK, He W, Nedergaard M. Locus coeruleus alpha-adrenergic-mediated activation of cortical astrocytes in vivo. Cereb Cortex. 2008;18(12):2789-2795.
64. Corkrum M, Covelo A, Lines J, et al. Dopamine-evoked synaptic regulation in the nucleus accumbens requires astrocyte activity. Neuron. 2020;105(6):1036-1047.
65. Takata N, Mishima T, Hisatsune C, et al. Astrocyte calcium signaling transforms cholinergic modulation to cortical plasticity in vivo. J Neurosci. 2011;31(49):18155-18165.
66. Picciotto MR, Higley MJ, Mineur YS. Acetylcholine as a neuromodulator: cholinergic signaling shapes nervous system function and behavior. Neuron. 2012;76(1):116-129.
67. Agarwal A, Wu PH, Hughes EG, et al. Transient opening of the mitochondrial permeability transition pore induces microdomain calcium transients in astrocyte processes. Neuron. 2017;93(3):587-605.
68. Kanemaru K, Sekiya H, Xu M, et al. In vivo visualization of subtle, transient, and local activity of astrocytes using an ultrasensitive Ca^{2+} indicator. Cell Rep. 2014;8(1):311-318.

69. Corkrum M, Rothwell PE, Thomas MJ, et al. Opioid-mediated astrocyte-neuron signaling in the nucleus accumbens. Cells. 2019;8(6):586.
70. Bengio Y, Lecun Y, Hinton G. Deep learning for AI. Commun ACM. 2021;64(7):58-65.
71. Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. Neuron. 2017;95(2):245-258.
72. Lake BM, Ullman TD, Tenenbaum JB, Gershman SJ. Building machines that learn and think like people. Behav Brain Sci. 2017;40:e253.
73. Marblestone AH, Wayne G, Kording KP. Toward an integration of deep learning and neuroscience. Front Comput Neurosci. 2016;10:94.
74. Wang Q, Ding SL, Li Y, et al. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell. 2020;181(4):936-953.e20.
75. Guo Fuglstad J, Saldanha P, Paglia J, Whitlock JR. Histological E-data Registration in rodent Brain Spaces. eLife. 2023;12:e83496.

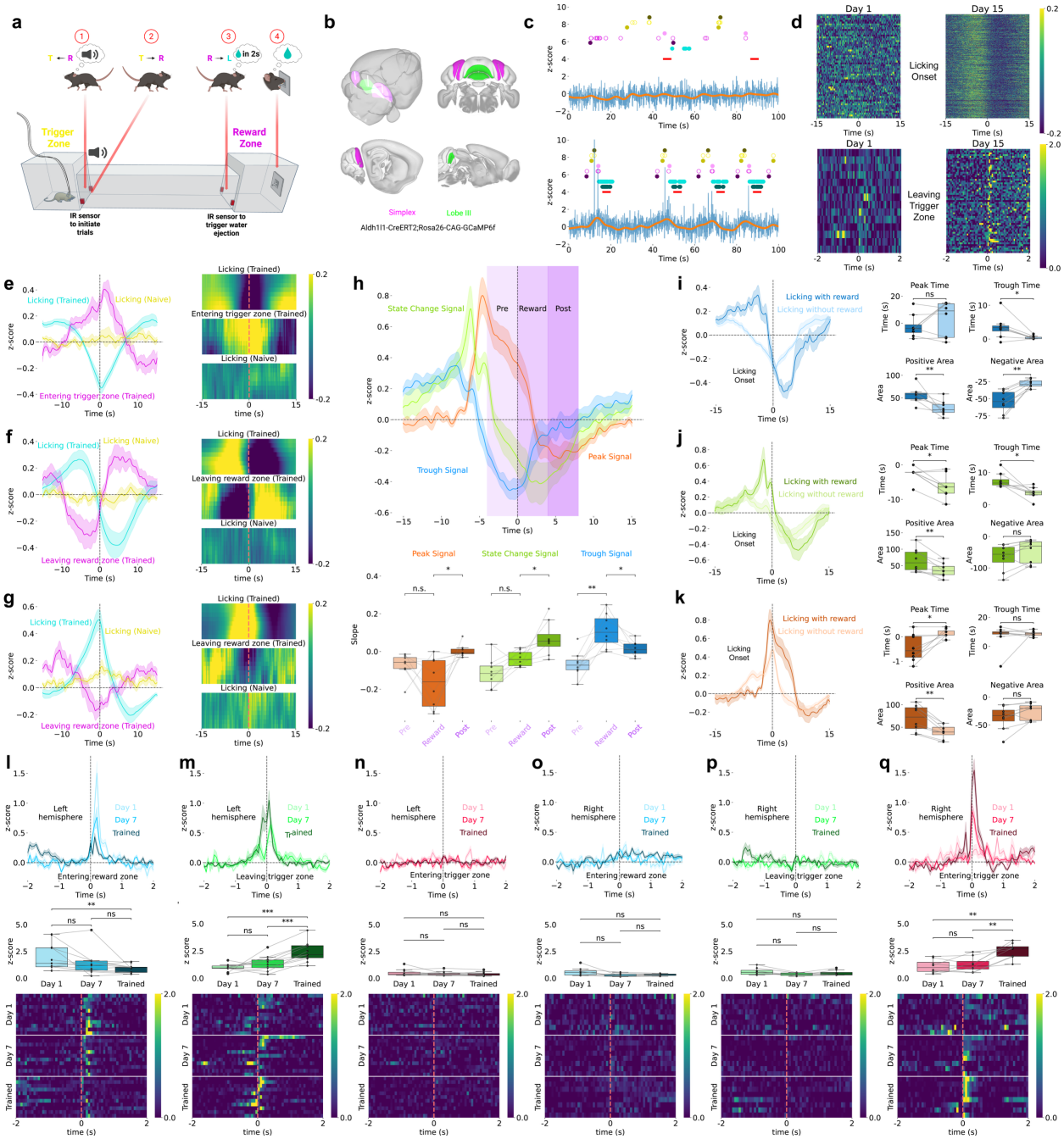


Figure 1: Cerebellar Astrocyte Ca^{2+} Dynamics During Reward-Guided Behavior

(a) Schematic of spatial sequence task showing four trial phases: entering trigger zone with auditory tone, leaving trigger zone toward reward zone, entering reward zone with 2-second delay, and licking during 4-second reward window. IR sensors monitor zone transitions and licking (BioRender).

(b) Fiber photometry recording sites in lobule III (green, medial vermis) and simplex (purple, lateral hemispheres) of Aldh1l1-CreERT2;Rosa26-CAG-GCaMP6f mice expressing GCaMP6f in Bergmann glia (HERBS, Allen Mouse Brain Atlas, and BioRender).

(c) Representative Ca^{2+} traces (z-scored ΔF/F) from Days 1 and 15. Blue: raw fluorescence. Orange: low-pass filtered (0.3 Hz) slow oscillations. Behavioral markers: reward zone IR breaks (purple), trigger zone IR breaks (yellow), licking (cyan), reward delivery (red). For zone IR breaks: hollow circles show all breaks; light solid circles show first break (entering); dark solid circles show last break (leaving). Day 15 shows robust task-aligned signals absent on Day 1.
(d) Heatmaps of activity (z-scored ΔF/F) aligned to licking onset (top) and leaving trigger zone (bottom), showing emergence of slow ramping signals and fast transients with training.
(e-g) Three slow signal types (z-scored ΔF/F; curve plots and heatmaps). **(e)** Trough signals (n=8 sites, 7 animals) decline before licking and peak at trigger zone entry. **(f)** State-change signals (n=8 sites, 6 animals) reverse polarity across reward transitions. **(g)** Peak signals (n=8 sites, 5 animals) ramp upward before licking and decline toward reward zone exit.
(h) Slow signal types (z-scored ΔF/F) aligned to reward onset with slope quantification across pre-reward (−4 to 0 s), reward (0 to 4 s), and post-reward (4 to 8 s) epochs.
(i-k) Rewarded versus unrewarded licking comparison (z-scored ΔF/F) showing differential dynamics in **(i)** trough, **(j)** state-change, and **(k)** peak signals.
(l-q) Fast Ca^{2+} transients (z-scored ΔF/F) show hemisphere-specific training-dependent changes. Left simplex (n=11 sites, 11 animals): **(l)** Entering reward zone response diminishes with training. **(m)** Leaving trigger zone response emerges with training. **(n)** Entering trigger zone/tone onset shows minimal response. Right simplex (n=8 sites, 8 animals): **(o)** Entering reward zone shows minimal modulation. **(p)** Leaving trigger zone shows no significant response. **(q)** Entering trigger zone/tone onset response increases with training.

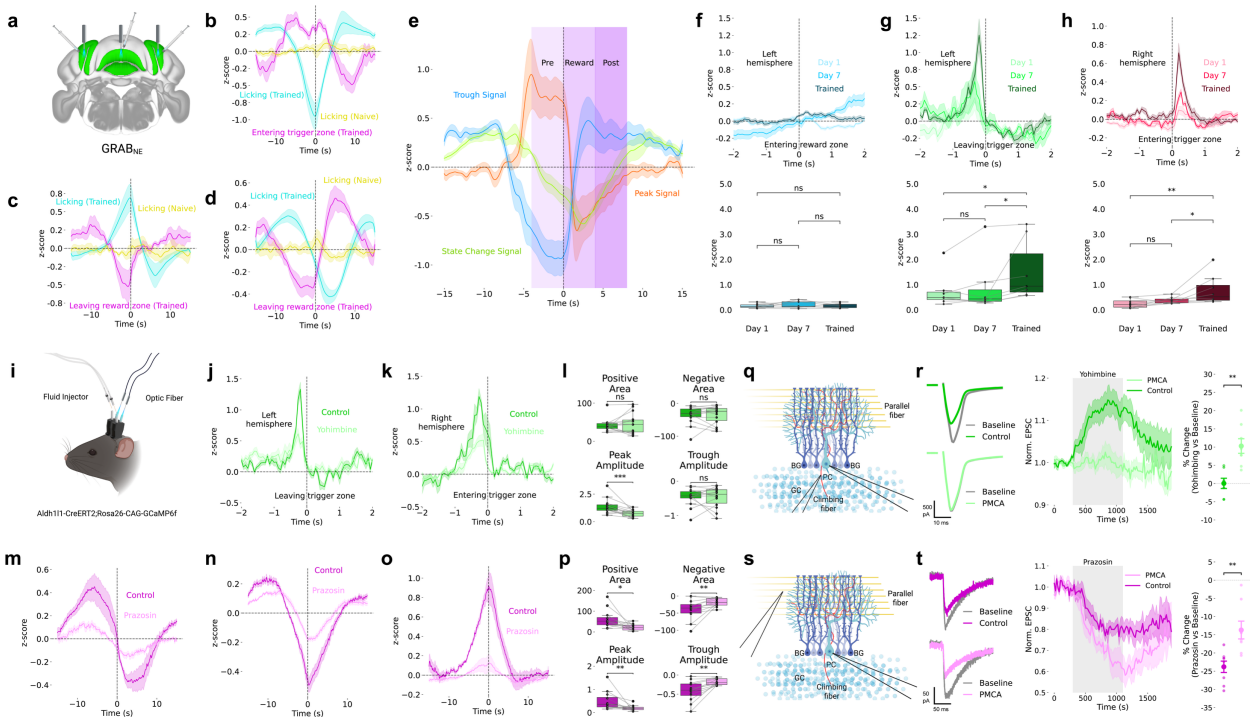


Figure 2: Norepinephrine-Dependent Modulation Through Distinct Adrenergic Receptors

(a) Schematic of GRAB_NE sensor expression in simplex and lobule III for measuring NE dynamics (HERBS, Allen Mouse Brain Atlas, and BioRender).

(b-d) Slow NE dynamics (z-scored ΔF/F) show three signal types: **(b)** trough (n=11 sites, 7 animals), **(c)** peak (n=7 sites, 6 animals), and **(d)** state-change (n=6 sites, 5 animals) patterns.
(e) Slow NE signals (z-scored ΔF/F) aligned to reward onset showing trough, peak, and state-change temporal profiles.
(f-h) Fast NE dynamics (z-scored ΔF/F) show hemisphere-specific training-dependent changes (n=8 sites per hemisphere). Left hemisphere: **(f)** minimal reward zone entry response, **(g)** emerging trigger zone exit response with training. **(h)** Right hemisphere shows progressive tone onset increases.
(i) Combined fiber photometry and microinfusion system schematic (BioRender).
(j-l) Yohimbine (α2-AR antagonist) suppresses fast Ca^{2+} transients (z-scored ΔF/F) in **(j)** left hemisphere leaving trigger zone and **(k)** right hemisphere entering trigger zone (n=14 sites, 8 animals).
(m-p) Prazosin (α1-AR antagonist) disrupts slow Ca^{2+} signals (z-scored ΔF/F): **(m)** state-change, **(n)** trough, and **(o)** peak signals show reduced amplitude (n=12 sites, 7 animals).
(q-r) Ex vivo climbing fiber EPSCs (n=8 control, n=7 PMCA cells, 5 animals each). Yohimbine enhances CF transmission; astrocytic PMCA reduces it.
(s-t) Ex vivo parallel fiber EPSCs (n=8 cells, 5 animals each). Prazosin and astrocytic PMCA reduce PF transmission.

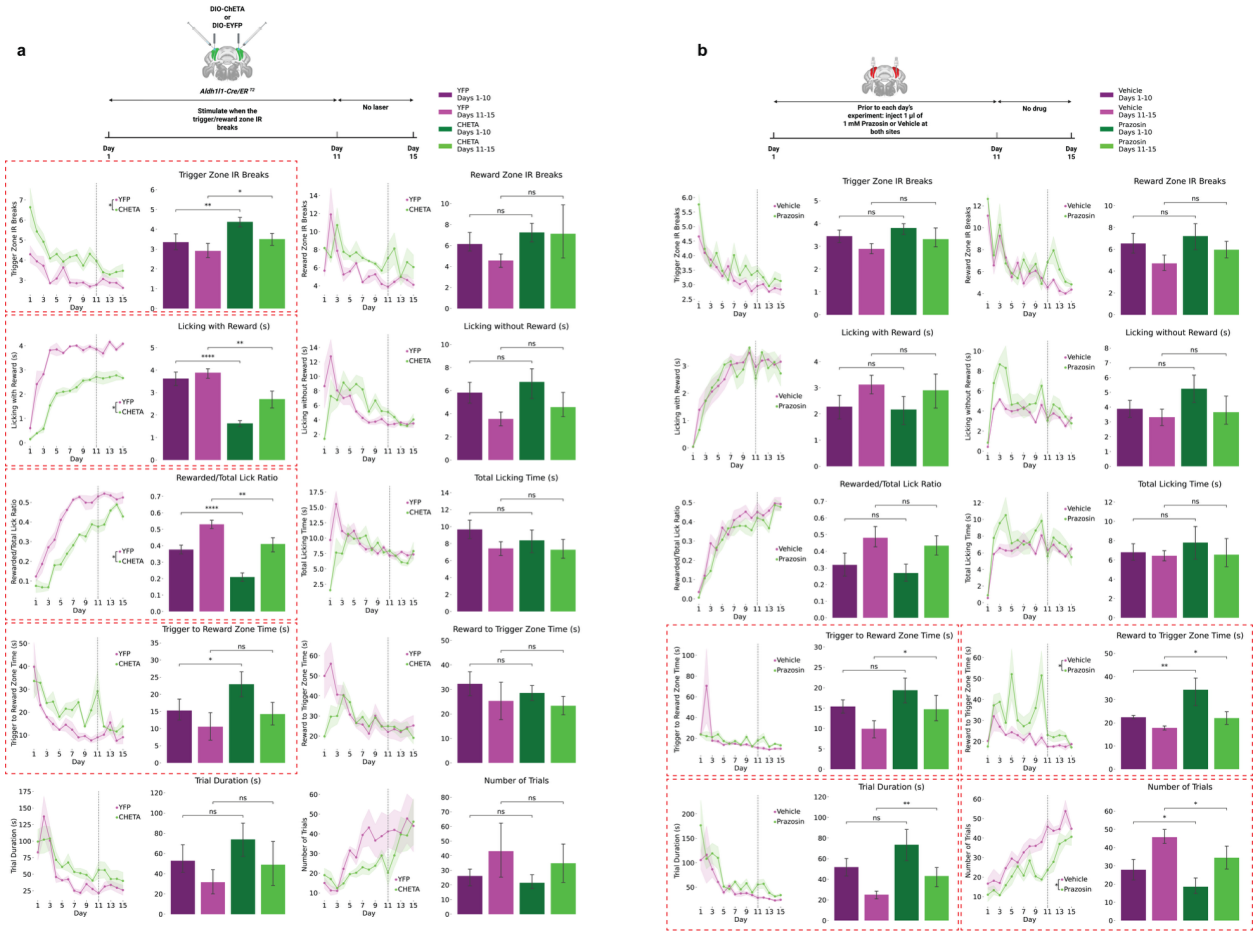


Figure 3: Dissociable Behavioral Roles of Fast and Slow Astrocytic Signals

(a) ChETA optogenetic stimulation of Bergmann glia during task events (n=6 ChETA, n=5 YFP control) impairs trigger zone IR breaks, licking with reward, rewarded/total lick ratio, and trigger-to-reward zone time.
(b) Prazosin infusion (n=6 prazosin, n=6 vehicle) increases trigger-to-reward and reward-to-trigger transit times, increases trial duration, and reduces trial number.

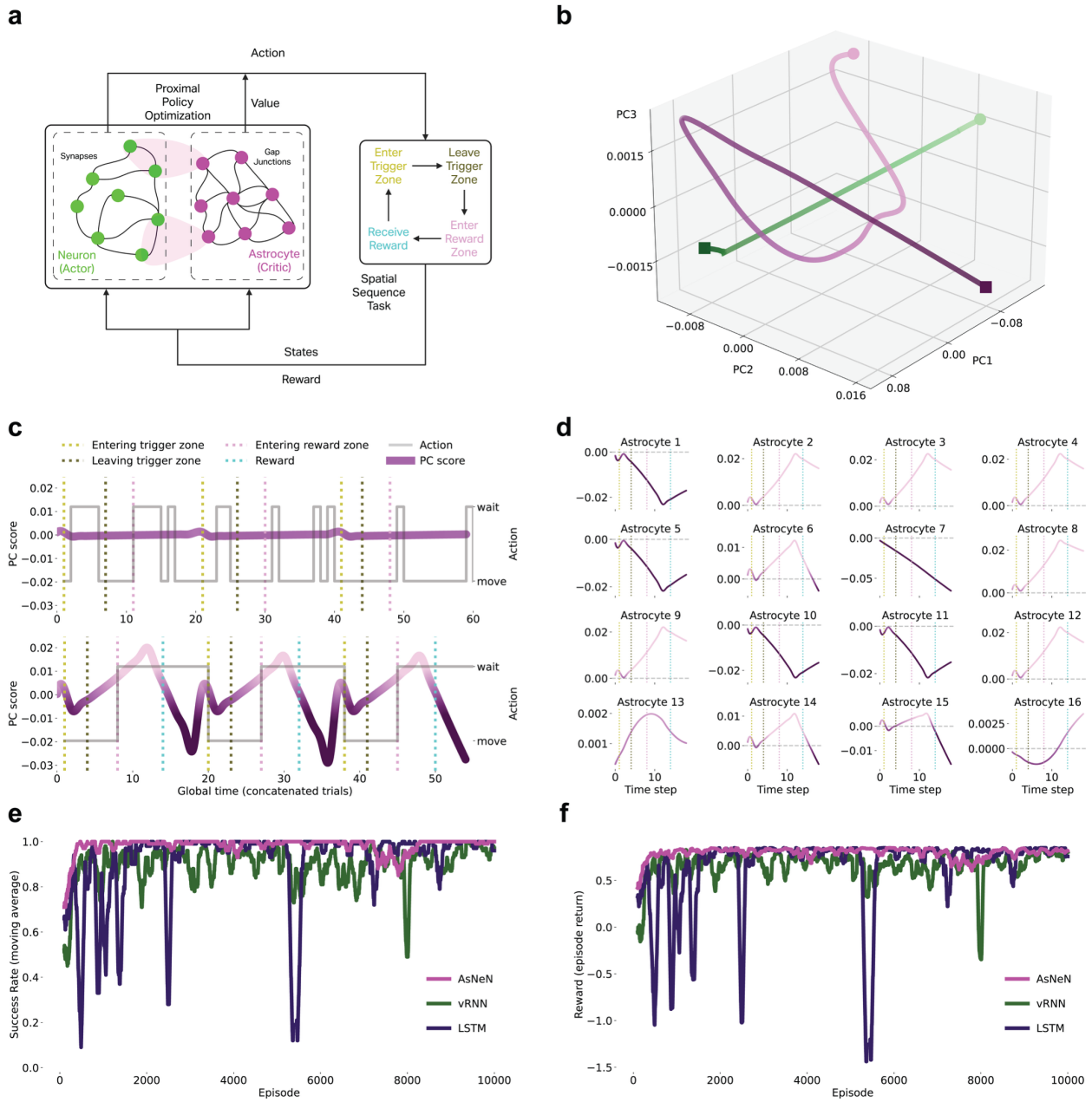
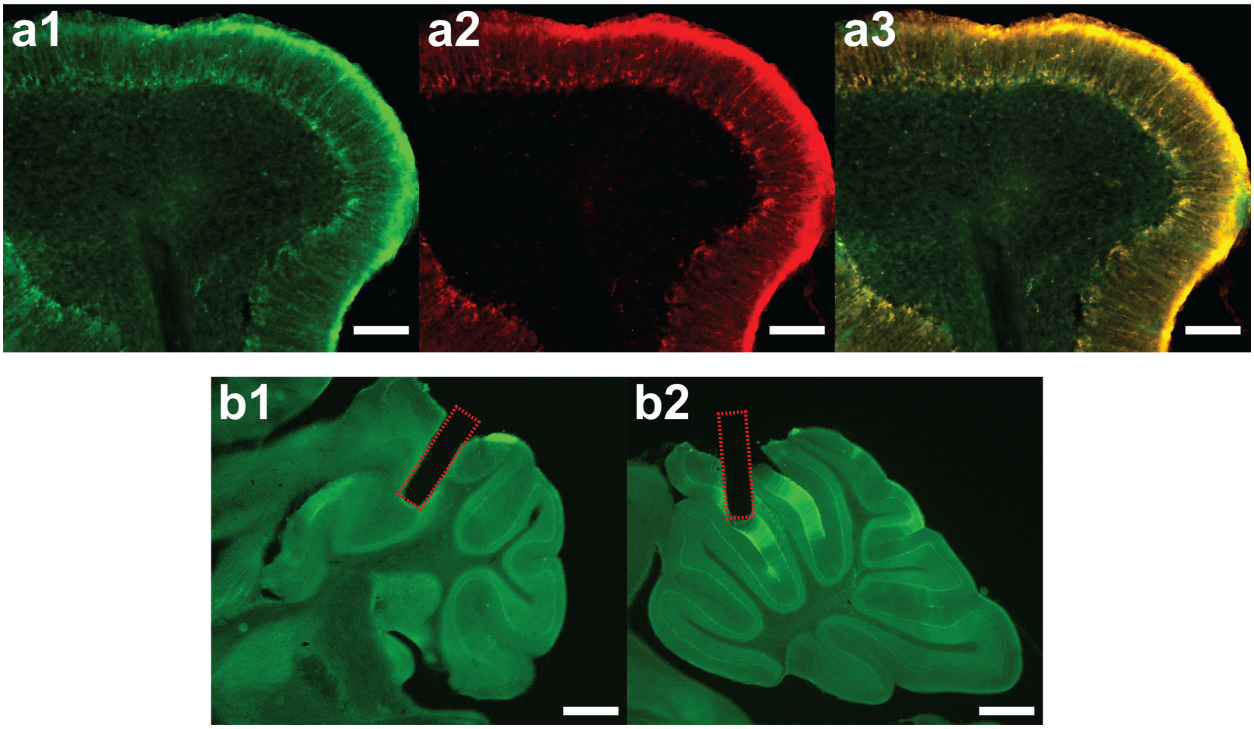
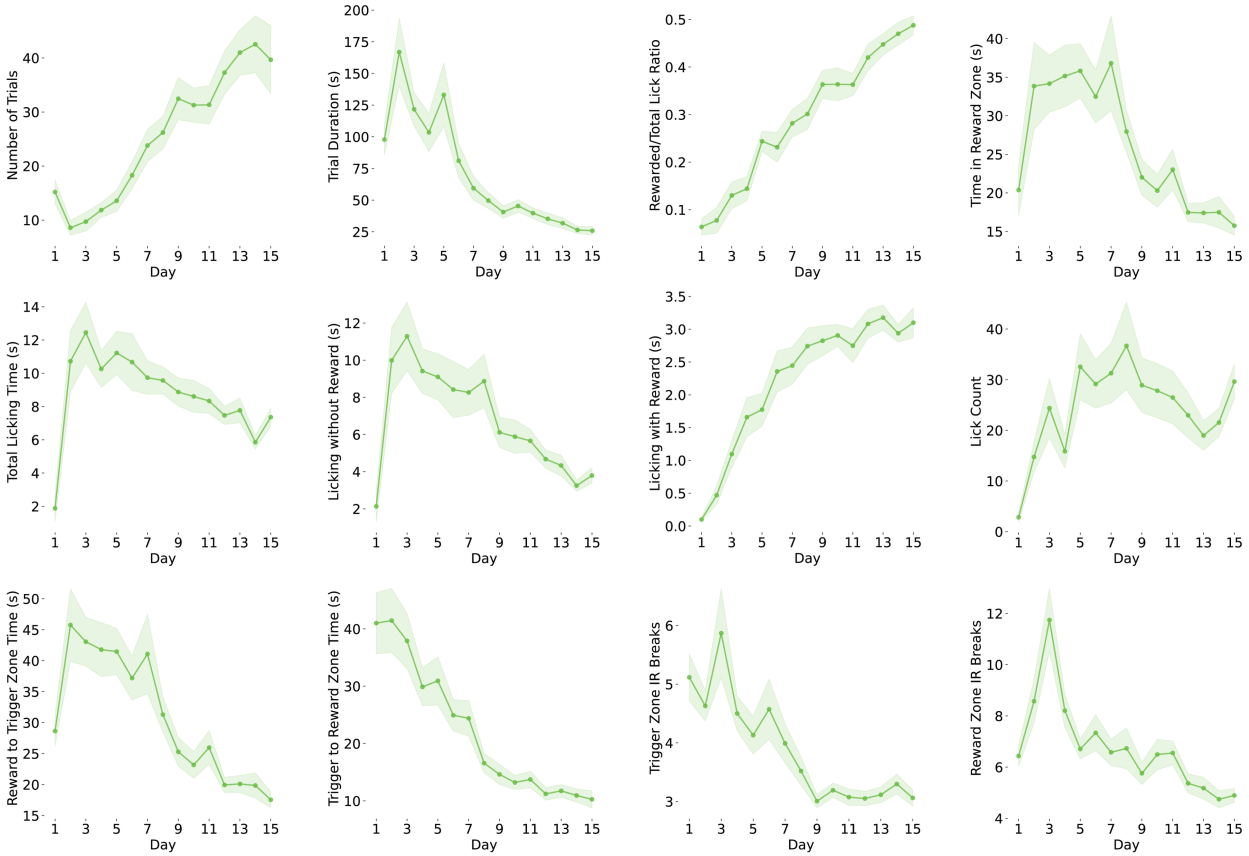


Figure 4: Actor-Critic Network Recapitulates Multitimescale Astrocytic Computation
(a) Biologically-inspired actor-critic network with 16 neuronal units (actor) and 16 astrocytic units (critic) implementing tripartite synapse principles. Network trained on spatial sequence task using PPO.

(b) PCA of neuronal (green) and astrocytic (magenta) activity showing distinct state-space trajectories.
(c) Population-averaged astrocytic signals pre- and post-training. Post-training shows fast transients at trigger zone entry and slow ramping signals with polarity reversals resembling biological state-change dynamics.
(d) Individual astrocyte response profiles across 16 units showing heterogeneous dynamics that recapitulate peak and trough signal diversity.
(e) Success rate comparison (moving average over 100 episodes) across training for AsNeN (neuron-astrocyte network), vRNN (vanilla recurrent neural networks), and LSTM (long short-term memory networks).
(f) Reward comparison (moving average over 100 episodes) across training for AsNeN, vRNN, and LSTM networks.

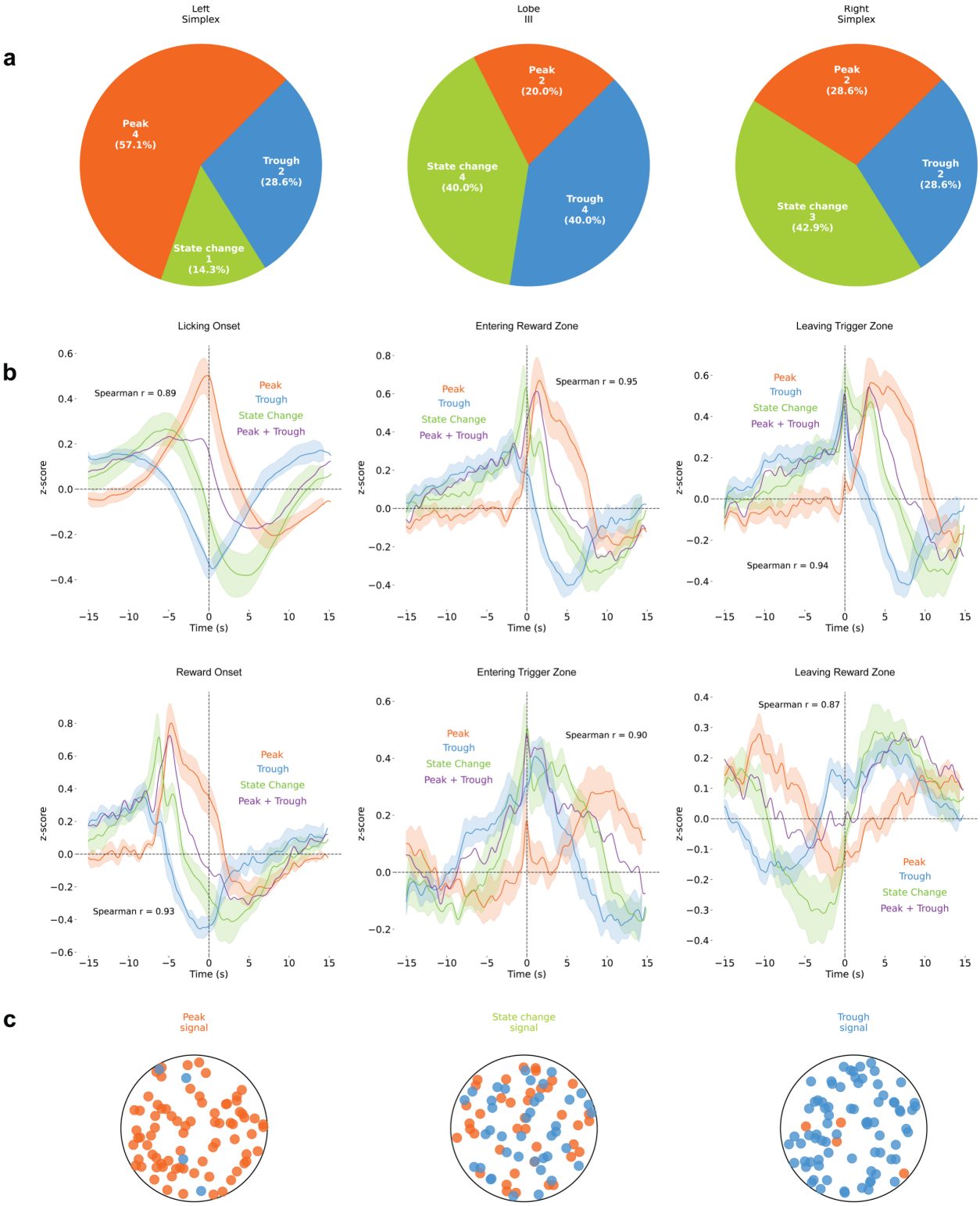


Supplemental Figure 1: Validation of Bergmann Glia Expression and Fiber Placement
(a1) GCaMP6f immunostaining (green) showing expression in Bergmann glia. Scale bar: 100 µm.
(a2) S100β astrocyte marker immunostaining (red). Scale bar: 100 µm.
(a3) Colocalization of GCaMP6f and S100β, confirming astrocyte-specific expression. Scale bar: 100 µm.
(b1) Fiber photometry cannula placement in left and right simplex. Scale bar: 1000 µm.
(b2) Fiber photometry cannula placement in midline lobule III. Scale bar: 1000 µm.



Supplemental Figure 2: Behavioral Learning Across Training

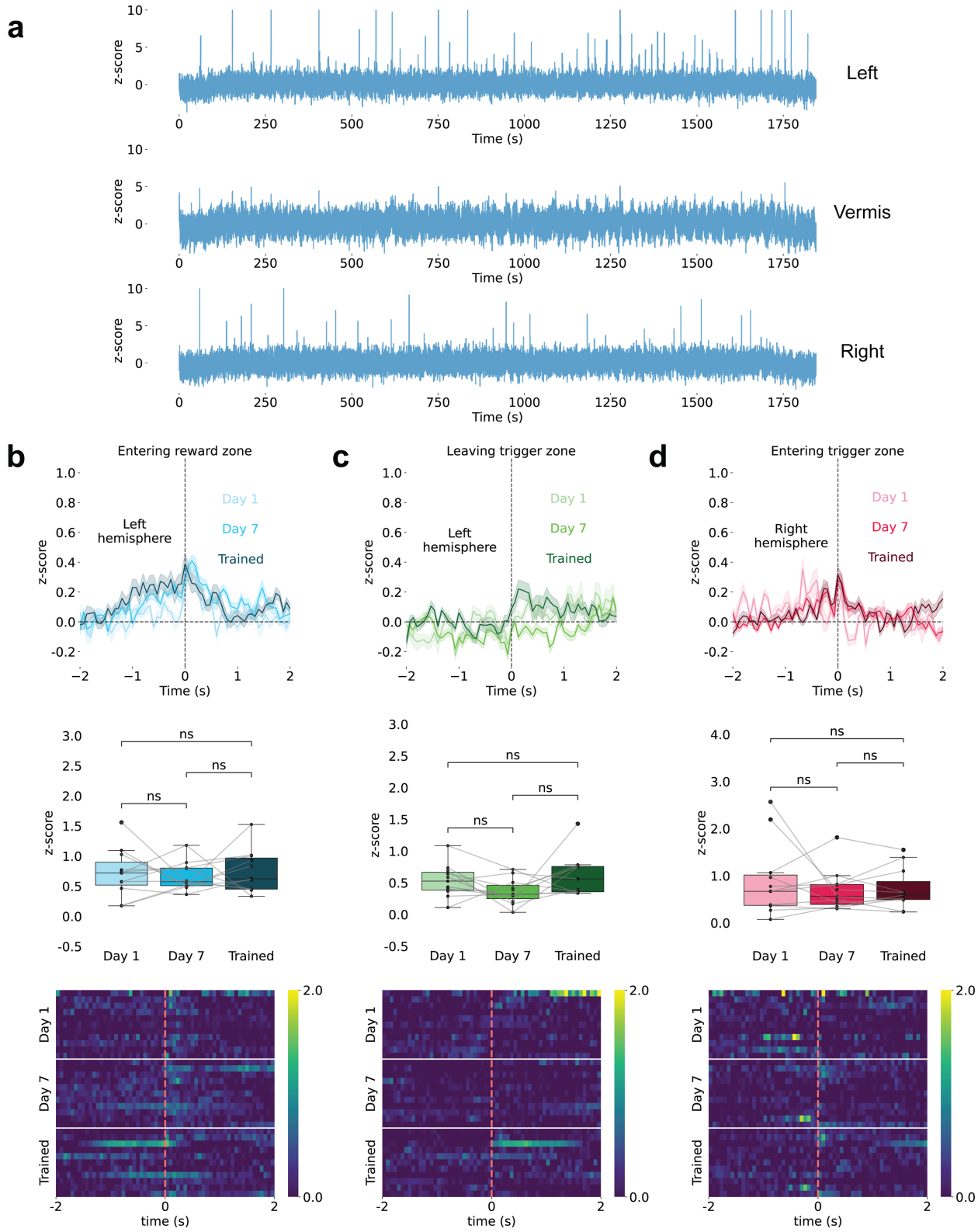
Training-dependent improvements (Days 1-15) in trial number, trial duration, rewarded/total lick ratio, reward zone time, total licking time, licking with/without reward, licking counts, trigger zone to reward zone time, reward zone to trigger zone time, and trigger zone/reward zone IR breaks.



Supplemental Figure 3: Spatial Distribution of Slow Signal Types Across Cerebellar Regions

(a) Signal type distribution across left simplex, lobule III, and right simplex shows no clear anatomical bias.

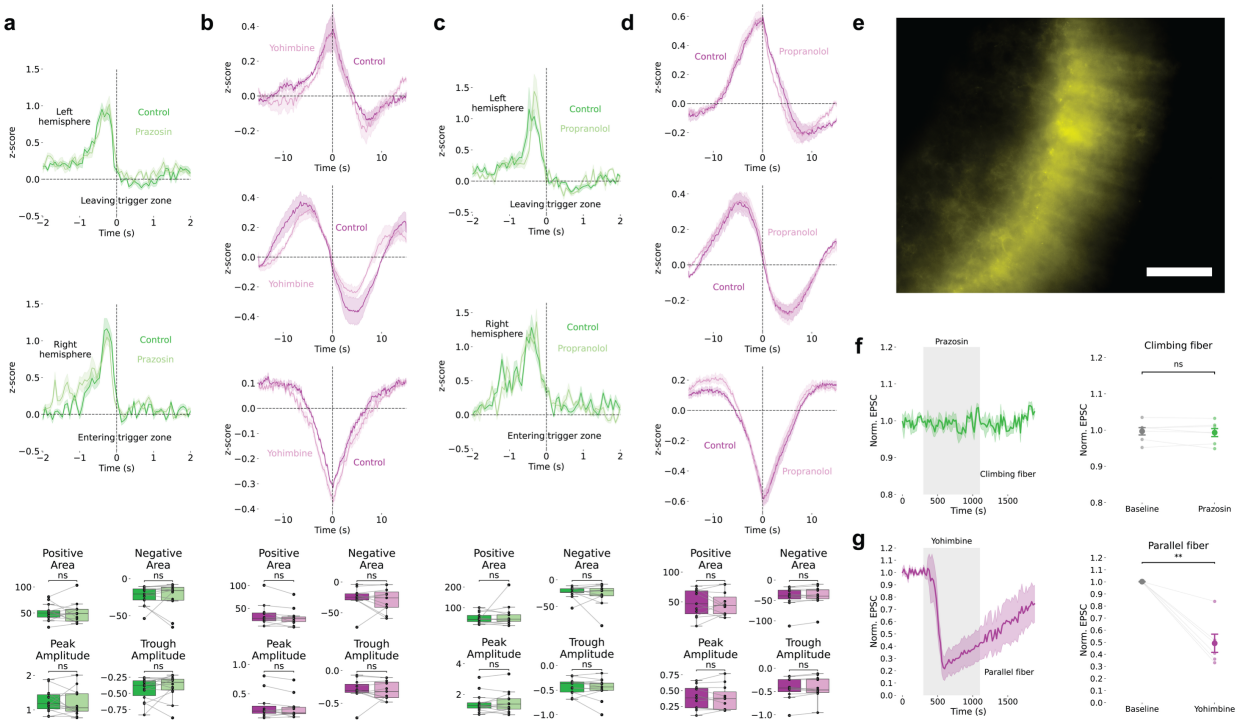
(b) Sum of peak and trough signals (z-scored ΔF/F) approximates state-change dynamics (Spearman r = 0.87-0.95), suggesting population integration.
(c) Schematic of heterogeneous signal types within photometry field of view.



Supplemental Figure 4: Raw Traces Showing Fast Transients Across Recording Regions

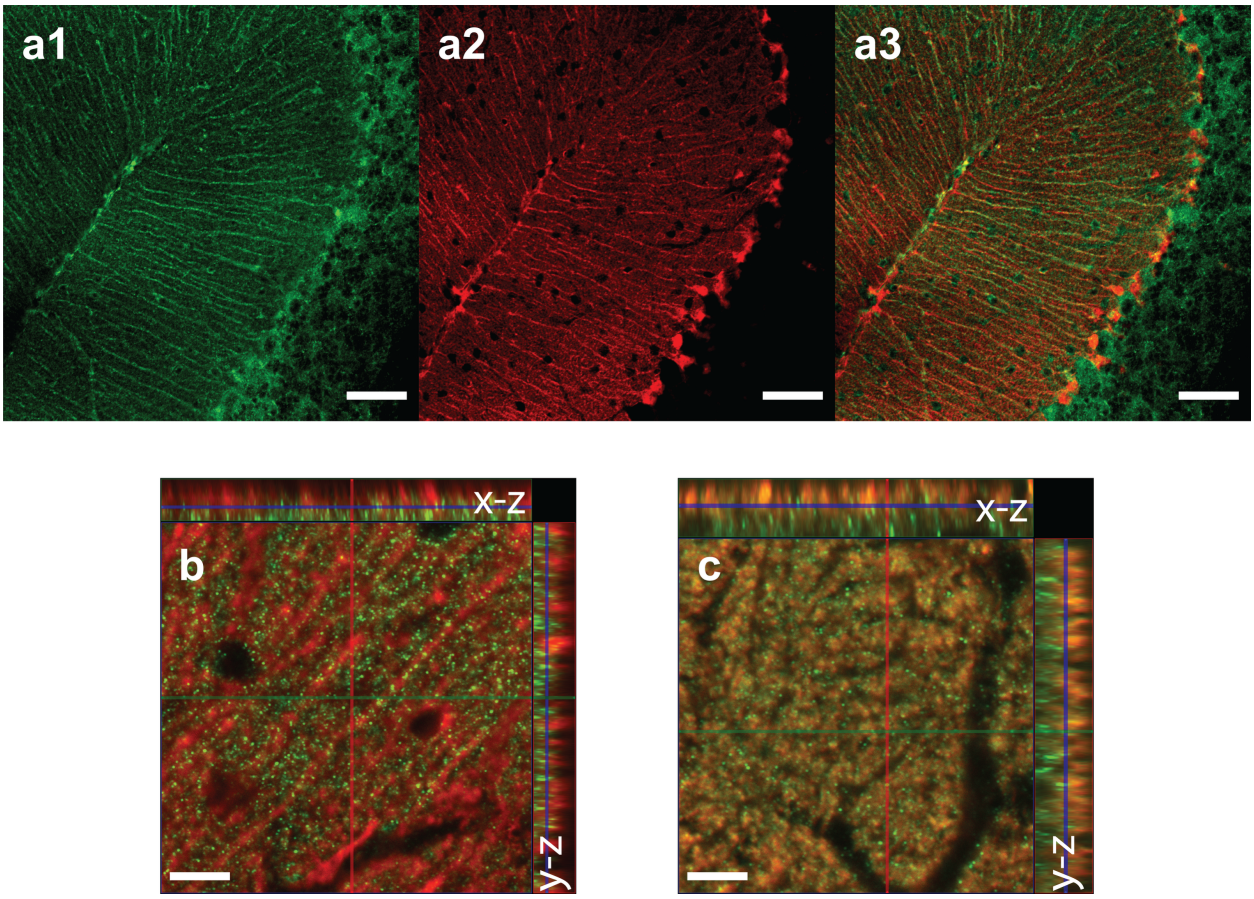
(a) Fast transients (z-scored ΔF/F) in lateral simplex but not midline lobule III.

(b-d) Lobule III shows no significant fast responses (z-scored ΔF/F) to **(b)** entering reward zone, **(c)** leaving trigger zone, or **(d)** entering trigger zone across training (n=11 sites, 11 animals).



Supplemental Figure 5: Pharmacological Specificity of Adrenergic Receptor Effects

(a) Prazosin (α1-AR antagonist) does not affect fast transients (z-scored ΔF/F).
(b) Yohimbine (α2-AR antagonist) does not affect slow signals (z-scored ΔF/F).
(c-d) Propranolol (β-AR antagonist) does not affect slow or fast signals (z-scored ΔF/F).
(e) PMCA expression in Bergmann glia. Scale bar: 50 µm.
(f) Prazosin does not affect CF-evoked EPSCs.
(g) Yohimbine reduces PF-evoked EPSCs.



Supplemental Figure 6: Anatomical Localization of Adrenergic Receptors

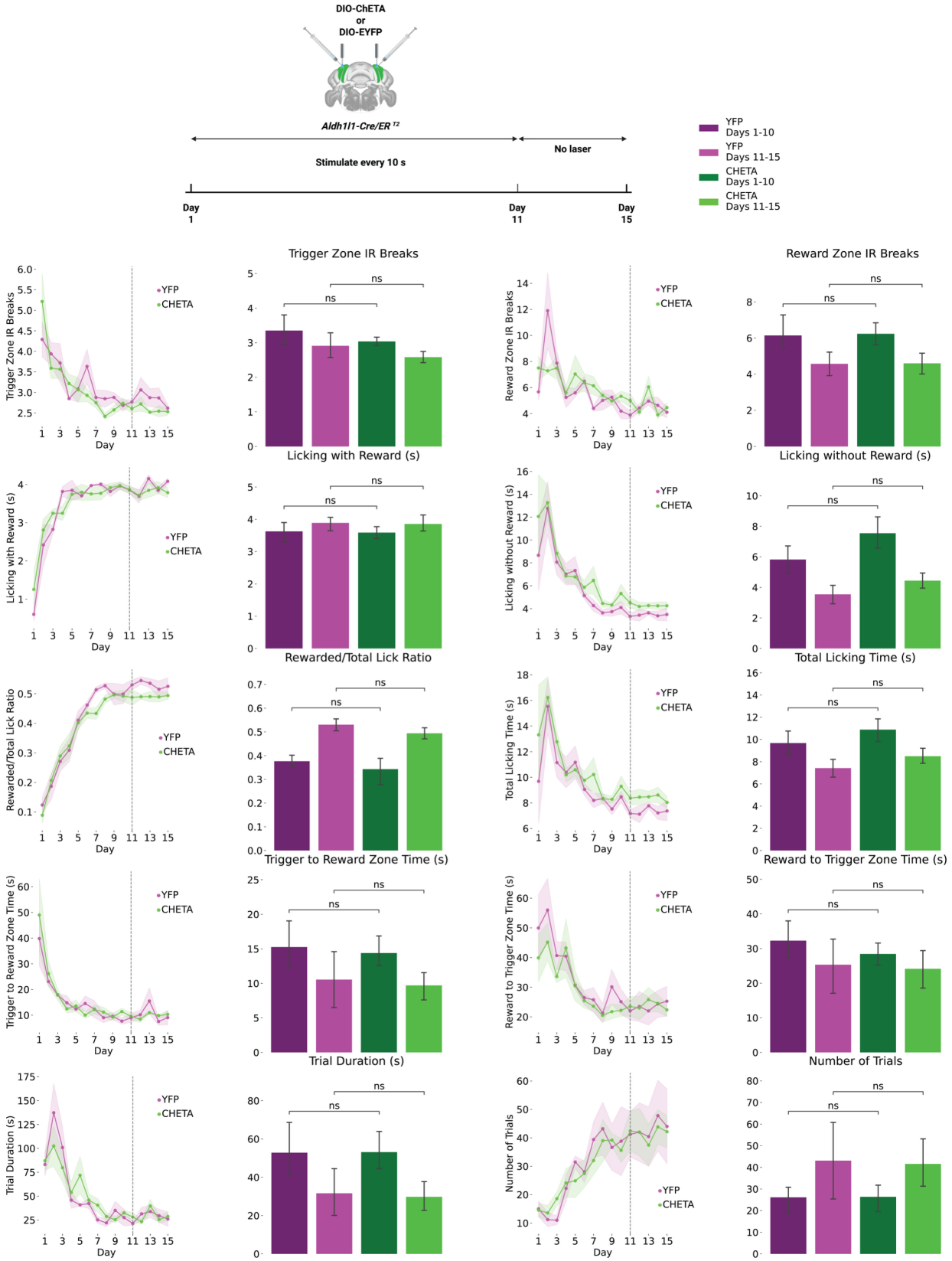
(a1) α2-AR immunostaining showing expression in cerebellar cortex Bergmann glia molecular layer. Scale bar: 50 µm.

(a2) S100β astrocyte marker immunostaining. Scale bar: 50 µm.

(a3) Colocalization of α2-AR and S100β. Scale bar: 50 µm.

(b) Colocalization of α1-AR and S100β astrocyte marker. Scale bar: 20 µm.

(c) Colocalization of α1-AR and VGLUT1-positive parallel fiber terminals. Scale bar: 20 µm.


DIO-ChETA
or
DIO-EYFP
Aldh1l1-Cre/ER T2
Stimulate every 10 s
No laser
Day 1
Day 11
Day 15
YFP Days 1-10
YFP Days 11-15
CHETA Days 1-10
CHETA Days 11-15
Trigger Zone IR Breaks
Reward Zone IR Breaks
Licking with Reward (s)
Licking without Reward (s)
Rewarded/Total Lick Ratio
Total Licking Time (s)
Trigger to Reward Zone Time (s)
Reward to Trigger Zone Time (s)
Trial Duration (s)
Number of Trials
YFP
CHETA
Day
ns

Supplemental Figure 7: Task-Independent Effects of Optogenetic Fast Signal Disruption

ChETA stimulation delivered at regular intervals (every 10 seconds independent of behavior) during Days 1-10 (n=7 ChETA, n=5 YFP control). Task-independent stimulation produces no behavioral changes compared to EYFP control, demonstrating that fast BG signal disruption requires temporal coupling to task-relevant cues (as shown in Figure 3) and is not due to nonspecific optogenetic activation.