Encoding without generation in the medial prefrontal cortex during cautious goal-directed actions

Muhammad S. Sajid*, Ji Zhou* and Manuel A. Castro-Alamancos

Department of Neuroscience
University of Connecticut School of Medicine
Farmington CT 06001 USA

* Equal contribution

Correspondence: Manuel Castro-Alamancos, PhD
Department of Neuroscience
University of Connecticut School of Medicine
Farmington, CT 06030-3401 USA
Email: mcastro@uchc.edu

Number of Figures: 12

Abbreviated title: Encoding vs causation in PFC

Acknowledgments: Supported by NIH grants to MAC.

Abstract

Adaptive behavior requires flexibly initiating or withholding actions in response to threat-predictive cues, often balancing urgency with cautious delay. Although dorsomedial, including prelimbic, prefrontal cortex (PFC) has been widely implicated in this control, the specific role of GABAergic neurons remains unclear. We combined fiber photometry and single-cell calcium imaging in mice performing cued tasks of increasing complexity to examine how PFC GABAergic neurons encode task variables, including relative to a comparison region, visual cortex (VI). Mixed-effects modeling showed that movement and baseline activity were dominant predictors of neural activity in both regions, but PFC neurons additionally encoded punished outcomes broadly and correct action generation or deferment only under challenging contexts. Single-cell analyses revealed subsets selectively modulated at cue onset or action generation. Yet broad optogenetic inhibition of PFC or VI produced similarly minimal effects on learning or performance across tasks, indicating that PFC encoding does not reflect behavioral control.

Significance statement:

Avoiding danger often requires deciding when to act and when to hold back. The medial prefrontal cortex (PFC) is widely assumed to support this control, yet the contributions of its neurons have remained unclear. Using neural population and single-neuron recordings together with targeted inhibition in mice performing learned threat-guided tasks, we show that GABAergic PFC neurons are activated by cues, actions, and outcomes, but are not required for executing the behavior itself. These findings suggest that the PFC primarily evaluates and contextualizes threat-motivated actions rather than generating them, highlighting a common mismatch between neural encoding and causal necessity.

Introduction

Adaptive behavior requires flexible selection and inhibition of actions in response to changing environmental contingencies, including cues that predict rewards or threats. Animals often must decide whether to act or withhold action to prevent harm signaled by learned cues (Thorndike, 1898; Mowrer, 1939; Bolles, 1970), such as a person actively crossing the street to evade a foe or passively waiting for a green light before crossing. Initiating an action promptly to prevent harm, yet not so prematurely that it incurs risk, defines cautious behavior, such as vacillating before crossing the street (Smith and Ratcliff, 2004; Bogacz et al., 2010; van Maanen et al., 2011; Guitart-Masip et al., 2012; Heitz and Schall, 2012; Ratcliff and Frank, 2012; Yee et al., 2022; Zhou et al., 2022). The dorsomedial, including prelimbic, prefrontal cortex (PFC) is widely thought to link predictive cues to appropriate actions while suppressing inappropriate or ill-timed ones (Ridderinkhof et al., 2004; Euston et al., 2012; Le Merre et al., 2021; Friedman and Robbins, 2022). Here, using calcium imaging, we examined how GABAergic neurons, which are key components of PFC circuit operations (Kawaguchi and Kubota, 1997; Gupta et al., 2000), encode threat-predictive and non-predictive cues, as well as the associated actions, in freely behaving mice that must either act or withhold action to prevent harm. Yet, despite their value, such studies raise two important concerns.

A major challenge in attributing neural activity to cue driven actions is that a large fraction of brain activity at any moment is tightly coupled to ongoing movement (Musall et al., 2019; Steinmetz et al., 2019; Stringer et al., 2019; Zagha et al., 2022; Hormigo et al., 2023a), a dominant behavioral-state category long-known to shape neural activation patterns across regions (e.g., (Castro-Alamancos and Connors, 1996; Castro-Alamancos, 2004a, 2004b)). This pervasive movement and state activation can confound interpretations of neural signals, leading

to misattributions of activity to task-specific variables such as cue encoding, action preparation or execution, and outcome processing. Studies examining PFC function during goal-directed actions under threat have not systematically monitored movement or dissociated its contribution from other predictors.

A second concern is that identifying neurons that encode task variables does not establish whether such activity is necessary for generating the behavior, since relations between neural activity and action are complex, often reflecting factors other than causal control (e.g., (Jonas and Kording, 2017; Krakauer et al., 2017; Shenoy and Kao, 2021; Barack et al., 2022)). To address both the encoding and necessity of PFC in cued, goal-directed behaviors, such as acting or passively deferring action to prevent harm, we used a two-pronged approach: (1) fiber photometry and single-cell calcium imaging of GABAergic PFC neurons in Vgat-Cre mice combined with mixed-effects modeling to control for movement contributions, and (2) optogenetic manipulations to test necessity. These approaches revealed that much of the apparent task-related activity in PFC could be explained by movement covariation or by sensory responses that were similar for threat-predictive and task-irrelevant cues. These sources of activity were also present in the visual cortex (VI), which is presumably less engaged in processing the auditory cues and threat behaviors we studied. Although GABAergic PFC neurons encoded punished errors and certain task contingencies in challenging contexts, broad optogenetic inactivation of PFC using different validated methods produced minimal behavioral effects across contingencies that were similar to VI inactivation. These findings underscore the dissociation between encoding and necessity and indicate that PFC contributes primarily to evaluative or contextual processing rather than directly generating cued, goal-directed actions under threat.

Results

PFC GABAergic activity is sensitive to movement

To assess the population activity of GABAergic neurons in PFC during spontaneous movement, we expressed GCaMP7f (Chen et al., 2013) via local injection of a Cre-dependent AAV in Vgat-Cre mice (n = 7) (Vong et al., 2011). A single optical fiber was implanted in the dorsomedial, *aka* prelimbic, area of the PFC (Mark et al., 2018), and calcium signals were recorded using fiber photometry, as previously described (Hormigo et al., 2021, 2023b). Figure 1A (left) shows a representative trajectory of an optical fiber targeting GCaMP-expressing GABAergic neurons in PFC. The estimated imaged volume extended ~200 µm from the fiber tip, encompassing approximately 2.5×10^7 μm^3 (Pisanello et al., 2019). As a control, we performed the same procedure in the visual cortex (VI), a functionally distinct cortical area. GCaMP7f was expressed in GABAergic neurons in Vgat-Cre mice (n = 5), and an optical fiber was implanted in VI (Fig. 1A middle). Figure 1A (right) includes a schematic showing the lens endings for both PFC and VI implantations based on histological reconstructions.

In freely moving mice, we recorded continuous calcium fluorescence ($\Delta F/F$) and spontaneous movement from three markers on the head to derive translational and rotational components, as the animals explored an open arena. To examine the relationship between movement and GABAergic neuron activation in the PFC and VI, we identified spontaneous movement peaks and aligned the fluorescence traces to those events, following procedures established in our previous studies (Hormigo et al., 2023b; Zhou et al., 2023, 2024). Movement peaks were categorized into three groups. The first category included all detected peaks (Fig. 1B, black traces), which revealed that VI neurons exhibited stronger activation than PFC neurons in response to equivalent movement (1WayAnova $F(1,583)= 30.8$ $p<0.001$). The second category

captured movement onsets from immobility, defined as peaks with no movement detected in the preceding 3 seconds (Fig. 1B, red traces). Again, activation was greater in VI than in PFC (1WayAnova $F(1,583)= 10.82$ $p=0.0011$), and notably, VI neurons showed a sharper drop in during the period of movement cessation preceding onset (1WayAnova $F(1,583)= 212.4$ $p<0.0001$), suggesting that VI activity more closely tracks ongoing motion. The third category included peaks separated by at least 5 seconds (Fig. 1B, cyan traces), minimizing the impact of closely spaced movements and isolating activity related to increases from an ongoing baseline. This condition also revealed stronger VI modulation by movement relative to PFC (1WayAnova $F(1,583)= 42.92$ $p<0.0001$). In all three categories, GABAergic activity in both regions increased significantly around movement compared to baseline. However, the consistently greater magnitude of activation in VI suggests that GABAergic neurons in this region more robustly encode general movement-related signals than those in the PFC.

To explore these relationships further, we computed cross-correlations between movement and GABAergic activity in the PFC or VI (Fig. 1C). Movement was significantly more correlated with VI activity than with PFC activity. When movement was decomposed into translational and rotational components, both were more strongly correlated with VI activity than with PFC activity, with translational movement contributing most strongly to the overall correlation (Fig. 1C). We next performed a linear fit between $\Delta F/F$ and movement, integrating both signals over a 200 ms window. This analysis revealed a strong positive linear correlation between VI activity and both translational and rotational movement, with correlations consistently stronger in VI than in PFC (Fig. 1D). Shuffling one variable in each pair abolished the observed correlations (Fig. 1D), confirming that the relationships do not reflect chance alignment.

We next asked whether PFC and VI GABAergic activity during movement depends on the direction of head turns relative to the recording site. Figure 2A shows head turns in the contraversive (cyan) and ipsiversive (red) directions, aligned to the corresponding ΔF/F signals recorded from PFC and VI neurons across the three categories of movement events. To assess direction-dependent responses, we compared the area under the curve, peak amplitude, and peak timing of the ΔF/F signals for equivalent ipsiversive and contraversive turns.

In the PFC, neuronal activity did not differ between ipsiversive and contraversive turns across any condition, indicating that GABAergic neurons in this region do not encode head turn direction (1WayRMAnova: $F(1,378)= 0.2$ $p=0.65$ *all turns;* $F(1,378)= 0.12$ $p=0.7336$ *no turns 3 sec prior;* $F(1,378)= 2.11$ $p=0.14$ *turns per >5 sec*). In contrast, VI neuron activity showed modest direction selectivity. For all turns, contraversive movements elicited significantly greater ΔF/Fresponses (1WayRMAnova $F(1,206)= 11.93$ $p=0.0007$ *all turns*). This effect remained significant for isolated turns ($F(1,206)= 10.88$ $p=0.0011$ *turns per >5 sec*), but not for turns following a period of quiescence ($F(1,206)= 3.59$ $p=0.059$ *no turns 3 sec prior*), suggesting that direction selectivity emerges in VI more clearly during ongoing movement. Together, these results indicate that PFC GABAergic neurons are insensitive to head movement direction, whereas VI GABAergic neurons have a modest preference for contraversive turns, particularly during sustained or ongoing movement. These results indicate that GABAergic neurons in both PFC and VI are sensitive to movement.

Cue Driven Action Selection Tasks

To evaluate the activation of cortical GABAergic neurons during cued goal-directed behavior, mice performed three successive procedures—noUS, AA19, and AA39—in a shuttle

box (Fig. 3A) during calcium imaging fiber photometry of PFC and VI neurons. In noUS (5 sessions), mice (n=7) were presented with three auditory tone CSs—CS1 (8 kHz), CS2 (4 kHz), and CS3 (12 kHz)—each delivered randomly at low, medium, and high intensity levels (~65-90 dB). The intensities were chosen to increase the likelihood of task errors, as mice are very proficient at performing these species-specific behaviors. Low-intensity cues are harder to detect, whereas high-intensity cues are harder to discriminate because loud sounds produce broader cochlear activation, reducing frequency specificity. CSs occurred in random order during 7-s action windows under an *ignore* contingency (Ig), in which they predicted no outcome but were terminated if the mouse crossed compartments (*spurious action*). Each trial was followed by a random intertrial interval during which crossings (ITCs) had no consequence. Spurious actions during CSs and ITCs thus provided measures of baseline exploratory responding and CS neutrality. In AA19 (7 sessions), stimulus presentation was identical to noUS, except that CS1 now incorporated an *active avoidance contingency* (AA), in which it predicted an aversive unconditioned stimulus (US) at the end of the 7-s window unless the mouse crossed compartments (*active avoid*). Failure to avoid resulted in rapid escape responses triggered by the US (*AA errors*). CS2 and CS3 remained non-predictive, as in noUS sessions. In AA39 (7 sessions), CS1 continued to signal that the animal must act to avoid the US, while CS2 now carried a *passive avoidance contingency* (PA); crossing during CS2 resulted in a brief (0.5 s) US (*PA errors*), which mice learned to avoid by withholding the action. CS3 remained unpredictive.

We analyzed behavioral and neural data using linear mixed-effects models with the following fixed factors: *task* (3 levels: noUS, AA19, AA39), *CS* (3 levels: CS1 [AA added in AA19], CS2 [PA added in AA39], CS3 [Ig]), *tone intensity* (3 dB levels: low, medium, high) and *group* (2 levels: PFC, VI), as well as nested random effects (sessions nested within mice). For

behavioral analyses, data from the PFC and VI groups are presented in combination after confirming no group effect ($\chi^2(1) = 0.95$, $p = 0.32$), while the other factors showed significant main effects and interactions with each other ($p < 0.0001$). Three behavioral variables were measured: actions, action latencies, and ITCs (Fig. 3B). Actions were defined as compartment crossings during CS presentations, though their meaning varied by task. In noUS, crossings were spurious actions to neutral cues. In AA19 and AA39, crossings during CS1 were correct active avoids, whereas failures to cross were active avoid errors (escapes). In AA39, crossings during CS2 were passive avoid errors, while withholding the action constituted a proper passive avoid. Thus, *Action* and *No-Action* trials were categorized as correct (active avoids, passive avoids and ignoring non-predictive cues) or errors (escapes, passive avoid errors, or spurious actions). Data from the same model, estimated across tone intensities, are shown in Fig. 3C.

During the initial noUS phase (Fig. 3B, open gray squares), mice showed no difference in the number of spurious actions across CS1–CS3 (Fig. 3C), indicating that the cues were initially neutral. Within each CS, however, tone intensity significantly modulated responding: high-intensity tones elicited more spurious actions than low-intensity tones ($t(2316) > 2.4$, $p < 0.05$). Notably, CS2 at high intensity produced significantly more crossings than CS2 at medium intensity ($t(2315) = 5.6$, $p < 0.0001$), suggesting that this tone was perceived as more salient.

In AA19 (Fig. 3B, red circles), mice made significantly more actions during CS1 compared to the noUS phase (Fig. 3C), indicating acquisition of active avoidance behavior to prevent the US. This increase was observed for all CS1 intensities (low: $t(1585) = 13.02$, $p < 0.0001$; medium: $t(1575) = 15.50$, $p < 0.0001$; high: $t(1585) = 16.89$, $p < 0.0001$). Spurious actions during CS2 and CS3 remained like noUS levels at low and medium intensities but rose significantly at high intensity (CS2: $t(1594) = 6.97$, $p < 0.0001$; CS3: $t(1585) = 3.93$, $p = 0.012$),

suggesting that discrimination between tones became more difficult as loudness increased. Within AA19, the number of active avoids during CS1 was markedly higher than spurious actions during CS2 ($t(2315) = 24.5$, $p < 0.0001$) or CS3 ($t(2315) = 25.4$, $p < 0.0001$), while CS2 and CS3 did not differ. Together, these results confirm robust and selective acquisition of active avoidance responding to CS1 and demonstrate that high intensity non-predictive cues generate discrimination errors.

In AA39 (Fig. 3B, cyan triangles), as in AA19, mice made significantly more actions during CS1 than during CS2 or CS3 (Fig. 3C; CS1 vs. CS2: $t(2316) = 21.45$, $p < 0.0001$; CS1 vs. CS3: $t(2316) = 28.6$, $p < 0.0001$). Compared to AA19, overall responding decreased across all CSs (CS1: $t(574) = 8.6$, $p < 0.0001$; CS2: $t(577) = 3.5$, $p < 0.0001$; CS3: $t(576) = 7.26$, $p < 0.0001$), indicating a general suppression of action when the passive avoidance contingency was introduced, consistent with increased task difficulty. Active avoids to CS1 declined at all tone intensities (low: $t(1599) = 6.07$, $p < 0.0001$; medium: $t(1599) = 5.43$, $p < 0.0001$; high: $t(1599) = 6.93$, $p < 0.0001$), as did spurious actions to CS3 (low: $t(1606.7) = 4.47$, $p = 0.0012$; medium: $t(1599) = 4.29$, $p = 0.0026$; high: $t(1607) = 6.79$, $p < 0.0001$). In contrast, CS2 actions—representing passive avoid errors—were selectively reduced at low intensity (low: $t(1611) = 3.4$, $p = 0.0018$; medium: $t(1603) = 2.09$, $p = 0.10$; high: $t(1611) = 1.9$, $p = 0.052$). Mice successfully inhibited actions to CS2 at low and medium intensities (<23% passive avoid errors), but at the high intensity they failed to withhold action on >50% of trials. Consequently, CS2 elicited more passive avoid errors than CS3 elicited spurious actions ($t(2315) = 7.13$, $p < 0.0001$), driven largely by the high-error rate to CS2-high. Together, these results indicate that introducing the passive avoidance contingency reduced overall action rates but that high-intensity tones impaired

response inhibition during CS2 trials likely because the loudness makes discrimination challenging.

We next analyzed action latencies, which can reflect increased caution or decision hesitation in challenging environments (Zhou et al., 2022). Action latencies did not differ between the noUS and AA19 phases for any CS (Fig. 3C). During AA39, however, latencies increased relative to AA19 for both CS1 active avoids ($t(1001) = 8.27$, $p < 0.0001$) and CS2 passive avoid errors ($t(1222) = 7.5$, $p < 0.0001$), but not for CS3 spurious actions ($t(1198) = 1.16$, $p = 0.68$). Considering tone intensity (Fig. 3B), CS1 avoid latencies increased at medium and high intensities (medium: $t(2160) = 5.51$, $p < 0.0001$; high: $t(2166) = 6.59$, $p < 0.0001$), but not at low intensity ($t(2170) = 3.15$, $p = 0.38$, n.s.). The lack of change at low intensity likely reflects already long latencies in AA19, consistent with reduced detectability of low-salience tones. Latencies of CS2 actions—representing passive avoid errors in AA39—also increased, particularly at low and medium intensities (low: $t(2222) = 5.53$, $p < 0.0001$; medium: $t(2210) = 4.34$, $p = 0.0042$; high: $t(2171) = 3.68$, $p = 0.061$). These elevated latencies suggest that mice hesitated longer before initiating actions under more challenging conditions that required inhibition, consistent with greater cognitive demand and response conflict during AA39.

Mice exhibited generally low ITC rates across all phases, indicating that CS-evoked actions occurred against a backdrop of minimal crossings outside of CS windows. Although ITCs were never punished (Zhou et al., 2022), their frequency decreased progressively as contingencies were added (Fig. 3C)—from noUS to AA19 ($t(271) = 2.78$, $p = 0.0057$) and from AA19 to AA39 ($t(271) = 3.34$, $p = 0.0018$)—reflecting reduced exploratory behavior as the task context became more demanding. When considering ITCs following specific CSs, rates did not differ among cues during noUS. In AA19, however, ITCs were fewer after CS1 trials than after

CS2 ($t(2315) = 8.4$, $p < 0.0001$) or CS3 ($t(2315) = 4.37$, $p < 0.0001$), consistent with successful active avoidance, whereas CS2 produced the highest ITC rates overall ($t(2315) = 4.1$, $p < 0.0001$ vs. CS3). A similar pattern emerged in AA39, except that ITC rates following CS2 and CS3 were comparable, both remaining higher than after CS1. Thus, ITC frequency tracked the behavioral contingencies associated with each cue and phase, reflecting decreasing exploration and greater behavioral control as task demands increased.

Movement dynamics during contingency changes

We next examined how movement dynamics evolved across task phases. Movement was quantified in three epochs relative to CS onset: baseline (pre-CS onset), orienting (0–0.5 s post-CS), and action (0.5–7 s post-CS). An additional from-action epoch (-2 to 2 s from crossing) captured movement surrounding the action event. The same mixed-effects model used for behavioral analyses was applied here, with an additional fixed factor, *Action*, separating trials indicating whether the animal crossed compartments during the CS window (2 levels: action, no-action). Figure 4A,B show baseline-corrected speed traces aligned to CS onset for action (A) and no-action (B) trials, while Figure 4C shows speed traces aligned from-action.

Baseline movement (pre-CS) varied systematically across task phases and action types (Fig 6A,B marginal means collapsed across intensities). There were strong main effects of Task ($\chi^2(2) = 52.0$, $p < 0.0001$) and Action ($\chi^2(1) = 199.9$, $p < 0.0001$), as well as significant CS × Action ($\chi^2(2) = 13.9$, $p = 0.001$) and Task × Action ($\chi^2(2) = 37.5$, $p < 0.0001$) interactions. During noUS, spurious actions were preceded by higher baseline speeds than no-action trials ($t(3052) > 7$, $p < 0.0001$ across all CSs), indicating that in the absence of reinforcement contingencies, spurious actions arose when mice were already active prior to CS onset. In AA19,

baseline speed remained higher for spurious actions compared to no-action trials for CS2 ($t(3082) = 4.44$, $p < 0.0001$) and CS3 ($t(3078) = 3.74$, $p = 0.00018$), but not between active avoids and escapes during CS1 ($t(3116) = 0.38$, $p = 0.7$), consistent with cued active avoids being stimulus-driven rather than determined by pre-existing locomotor state. In AA39, higher baseline movement continued to predict spurious actions (CS3: $t(3096) = 7.43$, $p < 0.0001$), passive avoid errors driven by CS2 ($t(3098) = 4.1$, $p < 0.0001$), and to a lesser degree active avoids (CS1: $t(3067) = 2.4$, $p = 0.015$), likely reflecting performance deterioration under the more challenging contingency. Across phases, baseline movement decreased from noUS to AA19 ($t(239) = 5.76$, $p < 0.0001$), consistent with the reduced intertrial crossings reported above, but did not further decline from AA19 to AA39 ($t(246) = 1.26$, $p = 0.2$), suggesting that baseline locomotor activity plateaued once the first avoidance contingency was established.

We then examined CS-evoked orienting responses (Fig. 6A,B), which reflect the rapid head and body movements immediately following CS onset (Fig. 4A,B) and are known to vary with both tone intensity and task-related behavioral state (Zhou et al., 2023). As expected, the strongest effect was a Task × dB interaction ($\chi^2(4) = 25.5$, $p < 0.0001$), with a modest dependency on CS (Task × dB × CS: $\chi^2(8) = 16.2$, $p = 0.04$). This pattern indicated that task-dependent changes in orienting magnitude were most prominent at the high tone intensity and for specific CSs. For instance, orienting responses to CS1 and CS3 increased between noUS and AA19 at high intensity (CS1-high: $t(3071) = 3.21$, $p = 0.0039$; CS3-high: $t(2861) = 3.4$, $p = 0.0019$), but not at lower intensities or for CS2. Because such intensity-dependent orienting effects could confound interpretation of subsequent movement during the action period, we defined the action window to exclude the orienting interval, thereby controlling for orienting-related variability in the mixed-effects model.

Movement during the action window was strongly influenced by CS ($\chi^2(2) = 25.7$, $p < 0.0001$) and tone intensity ($\chi^2(2) = 73.4$, $p < 0.0001$), with several significant interactions. The CS × Task interaction ($\chi^2(4) = 26.0$, $p < 0.0001$) indicated that the effect of each CS on movement varied across task phases, while the CS × dB interaction ($\chi^2(4) = 35.8$, $p < 0.0001$) revealed that these differences depended on cue intensity. As expected, movement was also strongly modulated by whether an action occurred ($\chi^2(1) = 1274$, $p < 0.0001$), and the magnitude of this effect varied across CS, Task, and tone intensity ($p < 0.0001$ for all interactions).

Movement during the from-action window, excluding the action factor, was strongly influenced by CS ($\chi^2(2) = 135.84$, $p < 0.0001$), task ($\chi^2(2) = 165.79$, $p < 0.0001$), and tone intensity ($\chi^2(2) = 39.89$, $p < 0.0001$), with several significant interactions. The CS × Task interaction ($\chi^2(4) = 259.42$, $p < 0.0001$) indicated that the effect of each CS on movement varied across task phases, and this somewhat depended on the tone intensity (CS × Task × dB; $\chi^2(8) = 18.05$, $p = 0.02$).

Focusing on Action trials from CS onset (Fig. 4A; Fig6 A,B), during AA19, active avoids elicited by CS1 produced greater movement than spurious actions during noUS ($t(1871) = 8.82$, $p < 0.0001$) across all three intensities. A similar, though weaker, increase was observed for CS2 ($t(2284) = 4.2$, $p < 0.0001$) and CS3 ($t(2048) = 3.31$, $p = 0.0028$), limited to medium and high intensities. Within AA19, CS1 active avoids evoked significantly greater movement than CS2 ($t(3102) = 3.14$, $p = 0.0033$) or CS3 ($t(3090) = 3.92$, $p = 0.00027$) spurious actions, whereas in noUS, movement did not differ among CSs when dB levels were pooled. When tone intensity was examined separately, CS2-high evoked more movement than the other high-intensity CSs during noUS, and in AA19 it no longer differed from CS1-high, consistent with its high salience.

Transitioning to AA39, movement associated with CS2 passive avoid errors was significantly elevated compared to CS2 spurious actions in AA19 ($t(2653) = 5.91$, $p < 0.0001$), particularly at medium and high intensities. In contrast, CS1 active avoids remained unchanged across intensities, whereas CS3 spurious actions decreased relative to AA19 ($t(2381) = 3.22$, $p = 0.0028$), mainly at high intensity. Within AA39, CS2 passive avoid errors were accompanied by greater movement than CS1 active avoids ($t(3136) = 5.6$, $p < 0.0001$), again most pronounced at medium and high intensities, likely reflecting movement triggered by the punishing consequences of CS2. Results aligned to CS onset were paralleled by analyses aligned from-action (Fig. 4C; Fig. 6A,B).

Importantly, in no-action trials, movement during correct CS2 passive avoids was reduced relative to CS2 no-action trials in the preceding AA19 phase ($t(1710) = 2.96$, $p = 0.0062$), with a similar reduction for CS1 no-action trials ($t(2408) = 2.63$, $p = 0.017$). No change was observed for CS3 no-action trials ($t(1466) = 0.23$, $p = 0.81$), consistent with the interpretation that CS1 and CS2 contingencies became behaviorally linked in AA39, where withholding action to CS2 and acting to CS1 represented opposing adaptive strategies.

Together, these results demonstrate that mice rapidly acquire distinct stimulus–action–outcome contingencies across progressively complex avoidance tasks. Behavioral performance, including avoidance rates and latencies, confirmed that animals flexibly adapted their actions based on cue valence and salience. Increasing task difficulty in AA39 impaired performance, reflecting reduced discriminability at high intensities and competing motivational drives. Movement analyses further revealed that learned avoidance actions (e.g., CS1 in AA19 and AA39) were accompanied by greater movement speed, whereas spurious actions during noUS were slower and largely reflected baseline movement. Overall, these findings validate this task

series as a robust framework for probing neural dynamics underlying stimulus-driven decision-making and motor output in the context of learned aversive outcomes.

PFC encodes errors

Next, we examined whether PFC GABAergic neuron activity encodes the behavioral contingencies manipulated across task phases and compared these activations to those of VI GABAergic neurons using fiber photometry. We applied the same mixed-effects modeling used in the behavioral and movement analyses, substituting ΔF/F as the dependent variable and adding Group (two levels: PFC, VI) as a fixed effect (note that Fig. 4 shows the speed traces, Fig. 5 shows the ΔF/F traces for PFC, and Fig. 6 shows the Models of speed (A-B), ΔF/F in PFC (C-D), and ΔF/F in VI (E-F)). Both regions exhibited robust CS- and task-related ΔF/F changes across noUS, AA19, and AA39 (Fig. 5 and Fig. 6C,D marginal means collapsed across intensities). Much of this activity was time-locked to orienting movements at CS onset and to subsequent avoid or escape actions, with stronger and sharper responses typically accompanying faster movements. These observations suggest that a substantial portion of the recorded neural activity was movement related. To separate contingency-related effects from those driven by movement or baseline activity, the models incorporated three covariates: window-specific head speed, baseline head speed, and baseline ΔF/F (the latter excluded in the baseline window model). Covariates were standardized within each window so that estimated marginal means reflected ΔF/F at average covariate values. For the action window, window-specific speed was standardized separately by outcome (action vs no-action), since movement magnitude differs substantially across outcomes (e.g., active avoids involve faster movement than passive avoids or escapes, which occur later after US onset). Thus, the action epoch model estimates neural

activity controlling for speed within action and no-action windows but not between them, whereas the from-action model enables comparing actions while controlling for movement. Random effects were sessions nested within mice. For each analysis window, we compared ΔF/F between action and no-action trials for each CS within a task to determine whether behavioral outcome predicted neural activity in PFC and VI regions (Fig. 6D,F). We also compared the activation for each CS between successive task phases to assess task-dependent changes in neural responses (Fig. 6D,E). Significant estimated marginal means contrasts are indicated in Figure 6, and relevant statistics are shown below.

During the baseline window, the mixed-effects model revealed a strong effect of baseline movement ($\chi^2(1) = 318.62$, $p < 0.0001$), underscoring the importance of controlling for movement when interpreting neural signals (Fig. 6C-F). After accounting for movement, we observed a significant Task × CS × Outcome interaction ($\chi^2(4) = 26.16$, $p < 0.0001$), indicating that baseline ΔF/F varied across task phases and behavioral outcomes. The four-way interaction including Group was only marginally significant ($\chi^2(4) = 8.18$, $p = 0.085$), suggesting that these effects depended only weakly on whether recordings were obtained from PFC or VI. Indeed, contrasts only revealed significantly higher baseline activation in PFC compared to VI neurons for CS3 during noUS spurious actions ($t(266) = 2.27$, $p = 0.024$) and no-action ($t(50) = 2.48$, $p = 0.016$) trials. Thus, while baseline neural activity varied with task contingencies and outcomes, this modulation was largely similar across cortical regions. To identify which specific CSs and outcomes contributed to these effects, we next examined contrasts within each group. In PFC neurons, baseline ΔF/F was modestly higher for action than no-action trials in CS1, with a more pronounced difference for CS3 in AA39. These effects suggest that in some cases, spurious actions were preceded by elevated baseline activity (Fig. 6D), even after accounting for

movement. Across tasks (Fig. 6C), only CS3 showed a significant increase in baseline activity for action trials from AA19 to AA39. For VI neurons, baseline ΔF/F was generally higher for action than no-action trials across several CSs and task phases (Fig. 6F), including spurious actions driven by CS1 in noUS, CS2 in AA19, and CS3 in both noUS and AA39. Notably, this relationship inverted for CS1 in AA19 ($t(3086) = 3.36$, $p = 0.00078$), indicating that higher baseline VI activity predicted escapes (errors) when the avoidance contingency was first introduced. Between tasks, baseline activity changed in opposite directions for active avoids and escapes in AA19, contributing to the divergence between these trial types. Together, these results reveal subtle regional differences, with VI neurons more sensitive to early contingency shifts predictive of error outcomes, whereas PFC neurons show more selective modulation preceding spurious actions. To control for baseline-related differences in neural activity and movement, subsequent models included both baseline movement and baseline ΔF/F as covariates.

During the orienting window (Fig. 5 and Fig. 6C-F), the mixed-effects model revealed strong effects of all three covariates—orienting movement ($\chi^2(1) = 96.09$, $p < 0.0001$), baseline movement ($\chi^2(1) = 44.08$, $p < 0.0001$), and baseline ΔF/F ($\chi^2(1) = 401.6$, $p < 0.0001$). After accounting for these covariates, we observed only a marginal Task × CS × Outcome interaction ($\chi^2(4) = 8.87$, $p = 0.061$), indicating that orienting ΔF/F varied only weakly across task phases and behavioral outcomes. The four-way interaction including Group was not significant ($\chi^2(4) = 5.05$, $p = 0.28$), suggesting that orienting activation patterns were comparable between PFC and VI. Indeed, contrasts only revealed a higher orienting activation in PFC compared to VI neurons for CS2 during A39 correct passive avoids ($t(33) = 2.64$, $p = 0.012$), suggesting an association between orienting activity and action inhibition in PFC. In PFC neurons, orienting-related ΔF/F was modestly higher for escapes than active avoids in CS1 and was elevated during spurious-

action relative to no-action trials in CS3 during AA19. Across tasks, orienting activity increased for CS2 no-action and CS3 spurious actions trials from noUS to AA19. Orienting activity also increased for CS1 escapes and CS2 action and no-action trials from AA19 to AA39, corresponding to the introduction of the more difficult contingency. In VI neurons, only CS3 during AA39 showed higher activation for spurious-action than no-action trials, and across tasks, only CS1 and CS3 exhibited significant increases in no-action trials from noUS to AA19. Together, these findings indicate that orienting-related activity showed limited modulation by task contingencies or outcomes and was broadly consistent across cortical areas. Nonetheless, within-group contrasts suggest that PFC neurons may exhibit slightly greater sensitivity to task transitions and error-related orienting responses than VI neurons.

During the action window (Fig. 5 and Fig. 6C-F), the mixed-effects model revealed strong effects of all three covariates—action-window movement ($\chi^2(1) = 306.78$, $p < 0.0001$), baseline movement ($\chi^2(1) = 196.08$, $p < 0.0001$), and baseline $\Delta F/F$ ($\chi^2(1) = 1874.7$, $p < 0.0001$). After accounting for movement within each outcome category (but not between outcomes), we observed a significant Task × CS × Outcome interaction ($\chi^2(4) = 10.16$, $p = 0.037$), indicating that action-related $\Delta F/F$ varied across task phases and behavioral outcomes. The four-way interaction including Group was also significant ($\chi^2(4) = 10.88$, $p = 0.027$), demonstrating that these effects differed between recordings obtained from PFC and VI. Contrasts revealed higher action-related activation in PFC compared to VI neurons for CS2 correct passive avoids ($t(23) = 3.65$, $p = 0.0013$) and CS1 escapes ($t(22) = 3.50$, $p = 0.002$) during A39, indicating PFC engagement during dynamic action inhibition, as these trial outcomes involve correct and incorrect action postponement during the action window, respectively. In PFC neurons, $\Delta F/F$ increased across tasks as active avoids and escapes emerged for CS1 between noUS and AA19.

For CS2, activity increased for no-action trials between noUS and AA19, and with development of both correct and error passive avoids from AA19 to AA39. Like CS1, CS3 showed increases for both spurious-action and no-action trials from noUS to AA19, indicating that the effects in CS1 may not specifically reflect active avoidance encoding but broader contextual changes.

In VI neurons, CS1 exhibited increases in activity only as escapes emerged between noUS and AA19, while CS2 showed increases with development of passive avoid errors from AA19 to AA39 likely reflecting the US delivery. CS3 displayed increases in no-action trials across task phases. Because movement was not controlled between outcomes in this window (by design), these differences (Action – No Action in Fig. 6D,F Action window) largely reflect ongoing movement-related activity and are deemphasized in the figure panel. Notably, these contrasts between action and no-action trials were stronger in VI than in PFC neurons, suggesting that VI GABAergic neurons more robustly encode movement during actions. Accordingly, direct comparisons of action-related neural activity are better evaluated in the from-action window, where movement effects are explicitly controlled.

During the from-action window (Fig. 5 and Fig. 6C-F), the mixed-effects model revealed strong effects of all three covariates—from-action movement ($\chi^2(1) = 154.73$, $p < 0.0001$), baseline movement ($\chi^2(1) = 40.467$, $p < 0.0001$), and baseline $\Delta F/F$ ($\chi^2(1) = 96.45$, $p < 0.0001$). After accounting for these covariates, we observed a significant Task × CS × Outcome interaction ($\chi^2(4) = 18.09$, $p = 0.0019$) and Task × CS × Group interaction ($\chi^2(4) = 22.65$, $p = 0.00015$), indicating that from-action $\Delta F/F$ varied across task phases, and behavioral outcomes or brain regions. The four-way interaction was also marginally significant ($\chi^2(4) = 9.6$, $p = 0.047$), demonstrating that the effects across tasks and outcomes differed somewhat between PFC and VI. Contrasts revealed higher from-action activation in VI compared to PFC neurons for CS2

($t(47) = 2.35$, $p = 0.022$) and CS3 ($t(38) = 3.2$, $p = 0.0027$) spurious actions in noUS, suggesting stronger movement related activation in VI prior to addition of reinforcement contingencies. In PFC neurons, $\Delta F/F$ around the time of action was higher for escapes than for active avoids in AA trials during both AA19 and AA39. Activity also increased for spurious actions in AA19 compared to noUS, but only for CS2, not for CS3. More prominently, CS2 activity increased for passive avoid errors in AA39 compared to spurious actions in AA19. These effects are consistent with enhanced encoding of error outcomes in PFC, such as escapes and passive avoid errors.

VI neurons showed little difference in the from-action window after controlling for movement. In contrast to the strong encoding of escapes in PFC across task phases, in VI this was only observed during AA39, not during AA19. There was also no enhancement of from-action activation when passive avoid errors emerged during AA39 compared to spurious actions in AA19, which was a prominent effect in PFC. Thus, VI GABAergic activity primarily reflects motor rather than contingency-specific components. Together, these findings indicate that when movement is controlled, PFC—but not VI—maintains sensitivity to contingency-specific neural activation at the time of action, encoding errors.

Across windows, PFC and VI neurons exhibited distinct patterns of activation reflecting different components of avoidance behavior. Baseline activity varied systematically with task contingencies and outcomes but was largely similar across cortical regions, indicating a shared modulation of context-related excitability. During the orienting window, both regions showed weak contingency-related modulation, with PFC exhibiting slightly greater sensitivity to task transitions. In the action window from CS onset, $\Delta F/F$ strongly tracked ongoing movement, particularly in VI neurons, highlighting their predominant motor encoding. Both PFC and VI enhanced their activation during passive avoid errors in AA39 compared to spurious actions in

AA19 when measured from CS during the action window. In addition, only PFC encoded correct passive avoids as they produced stronger activation than no action trials. Active avoids were associated with enhanced activation but this may reflect other factors as spurious actions showed similar effects, perhaps mice encode those spurious actions as avoids because they are not otherwise punished. During the from-action window, PFC neurons exhibited stronger activation to escapes than active avoids and to passive avoid errors than spurious actions across tasks, whereas these effects were weaker or largely absent in VI. Together, these results reveal a functional dissociation in which VI activity predominantly reflects motor output, while PFC GABAergic neurons encodes both success and failures of action generation and inhibition across active and passive avoidance contingencies, after controlling for baseline and movement-related factors.

PFC encodes avoidance in challenging contexts

To further interpret the results obtained in the full model across task phases, we next tested simpler models separately for AA19 and AA39 to ensure that phase-specific effects were evaluated independently and that variance or covariate structure from one phase did not influence estimates in another. Each model included Contingency (Ig and AA in AA19, with PA added in AA39; CS2 and CS3 were collapsed into Ig in AA19), Outcome, Group, and the covariates. This approach allowed a clearer assessment of how PFC and VI activity reflect contingency- and outcome-dependent modulation under easier (AA19) and more difficult (AA39) behavioral conditions.

In AA19 (Fig. 7A,B), the covariates retained strong effects across all windows, emphasizing the importance of controlling for movement and baseline activity. In the baseline

window, a significant Contingency × Outcome interaction ($\chi^2(1) = 12.82$, $p = 0.00034$) indicated that pre-CS activity predicted behavioral outcome differently across contingencies, but this relationship did not differ between PFC and VI, as no Contingency × Group or three-way interaction was observed. Specifically, in PFC, higher baseline activation preceding Ig trials was associated with spurious (error) actions, whereas in VI, higher baseline activation during AA trials was linked to errors (escapes). During the orienting window, there was a significant Outcome × Group interaction ($\chi^2(1) = 5.25$, $p = 0.022$), indicating that outcome-dependent modulation differed between regions. In PFC, orienting-related $\Delta F/F$ was elevated during spurious actions in Ig trials, whereas VI showed minimal outcome-related modulation.

In the action window (Fig. 7A,B), there were two-way interactions between the three factors but no three-way interaction. A Contingency × Group interaction ($\chi^2(1) = 9.25$, $p = 0.0023$) again indicated regional differences in contingency related activation. In PFC, escape responses during AA trials produced stronger activation than correctly ignored CSs in Ig trials, possibly reflecting late avoidance attempts that failed to reach the exit before the US onset. However, in both regions, spurious actions during Ig trials did not differ from active avoids in AA trials, indicating that neither PFC nor VI encoded active avoidance. In the from-action window, there was a significant Contingency × Outcome × Group interaction ($\chi^2(1) = 9.56$, $p = 0.0019$), demonstrating that outcome-dependent contingency effects varied by brain region. This reflected a selective enhancement of PFC activation during escapes relative to active avoids in AA trials, with no corresponding effect in VI. Importantly, there was no difference between spurious actions and active avoids in either region, indicating that neither PFC nor VI GABAergic populations encode active avoidance in AA19 when measured from the time of action or CS onset.

In AA39 (Fig. 7C,D), which introduced the PA contingency, the covariates retained strong effects across all windows, like in AA19. In the baseline window, as in AA19, there was a significant Contingency × Outcome interaction ($\chi^2(2) = 20.77$, $p < 0.0001$), indicating that pre-CS activity predicted behavioral outcomes differently across contingencies, but this relationship did not differ by brain region. Higher baseline activation preceding Ig trials was associated with spurious actions in both regions. Moreover, spurious actions showed higher baseline activation than active avoids in both PFC and VI, and compared to passive avoid errors but only in PFC.

During the orienting window (Fig. 7C,D), effects were stronger than in AA19, with significant Contingency × Outcome ($\chi^2(2) = 24.87$, $p < 0.0001$) and Contingency × Group interactions ($\chi^2(2) = 31.58$, $p < 0.0001$), but no three-way interaction. These results indicate that contingency-related modulation of orienting activity depends both on outcome and brain region, but not jointly. In PA trials, orienting-related $\Delta F/F$ was higher during passive avoid errors than correct passive avoids in both regions. In AA trials, escapes produced stronger orienting activation than active avoids, but only in PFC, whereas in VI, spurious actions during Ig trials elicited stronger orienting activation than correct ignores.

In the action window (Fig. 7C,D), all factors again showed two-way interactions, including a Contingency × Group effect ($\chi^2(2) = 24.14$, $p < 0.0001$), and additionally, a significant three-way Contingency × Outcome × Group interaction ($\chi^2(2) = 6.35$, $p = 0.041$), indicating that contingency- and outcome-dependent activation varied across regions. As in AA19, spurious actions during Ig trials did not differ from active avoids in either region, indicating that neither PFC nor VI encoded active avoidance when measured from CS onset. However, passive avoid errors produced stronger activation than active avoids and spurious actions in both regions.

Also, in PFC—but not VI—escapes during AA trials produced stronger activation than correct ignores in Ig trials and correct passive avoids in PA trials, indicating region-specific encoding of failed active avoidance, perhaps reflecting initiated but failed avoid actions as noted for AA19. More importantly, correct passive avoids produced stronger activation than correct ignores ($t(1301) = 4.14$, $p < 0.0001$), even though both contingency outcomes involve withholding action. This suggests that PFC specifically encodes the deliberate inhibition characteristic of passive avoidance, rather than the mere absence of action.

In the from-action window (Fig. 7C,D), as in AA19, there was a significant Contingency × Outcome × Group interaction ($\chi^2(2) = 11.78$, $p = 0.0027$), demonstrating that outcome-dependent contingency effects differed by brain region. This reflected selective enhancement of PFC activation during escapes relative to active avoids in AA trials, with no corresponding effect in VI. In both regions, passive avoid errors produced stronger activation than either active avoids or spurious actions, but these errors did not differ from escapes, suggesting that this effect may reflect the aversive US outcome, which is common to escapes and passive avoid errors. In contrast to AA19, and also to the action window during AA39, active avoids measured from the action produced stronger activation than spurious actions in PFC ($t(1362) = 2.52$, $p = 0.011$), but not in VI, indicating that PFC encodes active avoidance around the action time.

Together, these results indicate two main findings. First, PFC does not encode active avoidance when it emerges in simple contexts where decisions are easy and errors rare. In contrast, PFC encodes both passive and active avoidance under challenging conditions such as AA39, where stimulus discrimination determines whether action generation or inhibition is appropriate. Under these challenging contexts, cautious responding emerges and failures are frequent. The enhancement of PFC activity during active avoidance, relative to spurious actions,

appears only around the time of the action, perhaps reflecting encoding of cautious actions. PFC activity also increased during correct passive avoidance compared to correctly ignored cues, indicating engagement during action deferment. However, because PFC activity increased during both active and passive avoidance, this activation likely reflects elevated cognitive demands rather than encoding the specific action chosen. Second, error-related activation linked to punishment was present in both PFC and VI, but appeared more consistent and robust in PFC across contingencies, suggesting that GABAergic neurons in this region are particularly sensitive to aversive outcomes.

Error encoding in PFC reflects sensitivity to aversive, unsignaled footshocks

To probe the relation between aversive punishment and GABAergic neuron activation in both regions, we performed escape-only sessions in which the US, comprising white noise and footshock, was delivered without warning. These sessions included randomly interleaved trials of footshock alone, white noise alone, and the full US. During the baseline window, the mixed-effects model revealed a strong effect of baseline movement ($\chi^2(1) = 14.31$, $p = 0.00015$), underscoring the importance of controlling for movement. However, there was no effect of the US factor, the Group factor, or their interaction, indicating that the baseline activity did not differ by the different stimuli or brain region. During the escape window, from US onset (0 to 4 s), the mixed-effects model revealed effects of all three covariates—escape movement ($\chi^2(1) = 12.09$, $p = 0.00052$), baseline movement ($\chi^2(1) = 4.15$, $p = 0.041$), and baseline $\Delta F/F$ ($\chi^2(1) = 140.14$, $p < 0.0001$). After accounting for these covariates, we observed a significant main effect of US ($\chi^2(2) = 33.02$, $p < 0.0001$) and a US × Group interaction ($\chi^2(2) = 11.38$, $p = 0.0034$), indicating that orienting $\Delta F/F$ varied across the different stimuli and this depended on the brain region. In

particular, the foot-shock presented alone produced stronger activation than the full US ($t(105) = 3.89$, $p = 0.00034$) or the white noise alone ($t(105) = 6.13$, $p < 0.0001$, but this only occurred in PFC, not in VI (Fig. 7E,F).

The enhanced PFC response to footshock alone, despite the full US containing the same footshock, suggests that the presence of an auditory signal (white noise) diminishes the neural encoding of the aversive event. In contrast, the footshock alone may evoke stronger activation due to increased surprise, reflecting appraisal of unsignaled aversive stimulation. Notably, these effects were absent in VI, reinforcing the idea that PFC distinctly represents aversive stimulation.

GABAergic neuron class-specific encoding in PFC

We next used miniscopes to image activity from individual GABAergic PFC neurons (n=2217 neurons from 5 mice) while mice performed AA19 (Fig. 8; n=367) and AA39 (Fig. 9; n=1850). To classify the recorded neurons, we employed two steps. Initially, we calculated the cross-correlation between speed and $\Delta F/F$ for each neuron. Then, we applied k-means clustering to these cross-correlation functions to identify groups of neurons based on the lag and strength of the relationship between neural activity and movement. This analysis separated neurons into five classes (Fig. 8 and 9, inset “a”): neurons with strong correlations centered near zero lag (Class 5, cyan), three neuron classes with progressively weaker rightward-shifted correlations (Class 4. Blue; Class 2, red; Class 1, black), consistent with movement preceding neural activation to varying degrees, and neurons negatively correlated with movement (Class 3, green), indicating suppression during locomotion. The neuron Class was incorporated as a fixed-effects factor in the models and neuron ID was nested within mice in the random effects. Equivalence testing was used to confirm overlapping movement measures across neuron classes (Fig. 8 and 9 speed

traces), confirming that the Class factor effects did not reflect movement differences across task contingencies.

Figure 8A,B show the activity of these neuron classes and the associated movement during AA19 aligned to CS onset (A) and to action occurrence (B) for each contingency (Ig, AA) and outcome. The strongest activation was observed in Class 5 neurons (cyan, 6.8%), which showed peak cross-correlation near zero lag, followed by the Class 4 neurons (blue, 16.9%) whose activity closely tracked movement. Class 3 neurons (green, 15.8%) were inhibited around the time of action, whereas Class 1 (black, 27.3%) and Class2 (red, 33.2%) neurons were weakly activated. Notably, Class 5 neurons were also active during CS1, CS2, and CS3 no-action outcomes, when movement remained near baseline, suggesting that they may respond to the CS itself. Their strong activation at CS onset supports this interpretation.

In the mixed-effects models (Fig. 8C,D), the covariates retained strong effects across all windows. In the baseline window, a significant Contingency × Outcome interaction ($\chi^2(1) = 44.01$, $p < 0.0001$) and a three-way interaction with Class ($\chi^2(4) = 11.21$, $p = 0.024$) revealed distinct pre-CS modulations among neuron classes. Classes 1 and 2 showed lower baseline activation for correct active avoids than for spurious actions, while escapes had higher baseline activation than correctly ignored CSs in Classes 1, 2, and 4 (Fig. 8C). Thus, elevated baseline activation was associated with active avoidance errors across these three classes (Fig. 8D).

During the orienting window, there was a significant main effect of neuron Class ($\chi^2(4) = 124.33$, $p < 0.0001$) but the interaction with Contingency × Outcome was weak ($p = 0.08$), indicating that orienting-related $\Delta F/F$ differed mainly by neuron class. Only Class3 neurons showed higher activation during active avoids compared to spurious actions, suggesting that they selectively encoded avoidance (Fig. 8C). In Ig trials, orienting-related activation during spurious

actions was higher than during correctly ignored CSs in Classes 2 and 5, whereas Class3 neurons showed the opposite effect, consistent with their distinct (negative) cross-correlation with movement.

In the action window, there was a Contingency × Outcome interaction ($\chi^2(1) = 20.47$, $p < 0.0001$), but no three-way interaction with Class ($p = 0.08$). Escapes in AA trials produced stronger activation than correctly ignored CSs in all classes except Class 1. Class 3 neurons showed higher activation during active avoids than spurious actions, whereas Class 2 neurons showed the inverse pattern—consistent with their opposite cross-correlation polarity. Thus, as in the orienting window, only Class 3 neurons selectively encoded active avoidance with higher activation relative to spurious actions (Fig. 8C). Comparisons between outcomes (Fig. 8D) are deemphasized because movement was not controlled across outcomes by design; accordingly, these differences likely reflect movement-related activity, which was opposite in Class 3 neurons compared to the other classes—showing stronger activation for no-action (low movement) than for action (higher movement) trials.

In the from-action window, a significant Contingency × Outcome × Class interaction ($\chi^2(4) = 17.33$, $p = 0.0016$) demonstrated that action-related activity depended on neuron class. Only Class 5 neurons showed stronger activation for active avoids than for spurious actions, indicating that this subset of GABAergic neurons—strongly movement-correlated with minimal lag—selectively encodes the avoidance action itself. For most other neuron classes, escapes produced stronger activation than active avoids, consistent with the aversive stimulation effect observed in the photometry experiments. Notably, this escape-related enhancement was absent or variable in Class 3 and Class 5 neurons, which encode avoidance from CS onset or the avoidance action, respectively.

Together, these results indicate that during AA19 two neuron classes are closely linked to active avoidance behavior but in distinct ways. Class 3 neurons are inhibited by movement and show weak cue-evoked activation that nevertheless differentiates avoidance from spurious actions. Class 5 neurons, which are robustly activated by both movement and the cues, do not distinguish active avoids from spurious actions unless activity is aligned by the action itself. The remaining classes primarily reflect responses to the aversive stimulation.

PFC encoding shifts in challenging contexts

We next performed the same analysis for neurons recorded during AA39, a more complex situation. Although these recordings were obtained from the same animals, they reflect independent neuron samples. As during AA19, cross-correlation and clustering analyses revealed five classes of neurons with similar relation to movement (Fig. 9, a inset). Figure 9A,B show the activity of each neuron class during AA39 aligned to CS onset (A) and to action occurrence (B), together with the associated movement traces for each action type. In AA39, CS2 incorporated the PA contingency, and therefore action during CS2 signifies passive avoid errors, while no-action entails a proper passive avoidance.

The strongest activation was observed in Class 5 neurons (Fig. 9A,B; cyan, 10.9 %), which showed peak correlation near zero lag, followed by Class 4 neurons (blue, 10.3 %) whose responses closely tracked movement, and much more weakly Classes 2 (red, 22.7 %) and 1 (black, 32.4 %). Class 3 neurons (green, 23.7 %), which were inhibited in relation to movement, showed little activation during AA39.

In the mixed-effects models (Fig. 9C,D), the covariates retained strong effects across all windows. In the baseline window, a significant Contingency × Outcome interaction ($\chi^2(1)$ =

194.01, $p < 0.0001$) and a three-way interaction with Class ($\chi^2(4) = 66.83$, $p < 0.0001$) revealed distinct pre-CS modulations among neuron classes. Classes 1, and 2 showed lower baseline activation for active avoids and passive avoid errors than for spurious actions, while escapes had higher baseline activation than correctly ignored CSs or correct passive avoids in Classes 2, 3, 4, and 5 (Fig. 9C). All classes, except Class 3, showed elevated baseline activity preceding spurious actions in Ig trials. In AA trials, escapes were preceded by higher baseline activity in Classes 2, 3, and 4. In PA trials, errors were predicted by higher baseline activity in Class 1 neurons, which had the weakest relation to movement, whereas this relation was inverted for Class 2 neurons. Thus, elevated baseline activation was associated with a greater likelihood of spurious actions or AA errors across most neuron classes, but this relation was more complex in the PA contingency, where Class 1 and 2 neurons showed opposite patterns.

During the orienting window, there was a significant Contingency × Outcome interaction ($\chi^2(2) = 106.17$, $p < 0.0001$) and a strong three-way interaction with Class ($\chi^2(8) = 81.89$, $p < 0.0001$), indicating that orienting-related $\Delta F/F$ varied across contingencies and neuron classes. Across most classes, CSs predicting AA or PA contingencies evoked stronger orienting-related activation than those predicting Ig trials, irrespective of outcome (Fig. 9C). In Classes 1, 5, and to a lesser extent Class 3, AA and PA errors as well as spurious actions were accompanied by heightened CS-evoked activation, suggesting that elevated PFC activity at cue onset predicts behavioral outcomes.

In the action window, there was a Contingency × Outcome interaction ($\chi^2(2) = 179.6$, $p < 0.0001$), and a three-way interaction with Class ($\chi^2(8) = 240.7$, $p < 0.0001$). For all classes, escapes in AA trials produced stronger activation than correctly ignored Ig trials and correct PA trials, likely reflecting late, failed avoidance actions, as this window precedes the escape interval.

In Classes 2, 4, and 5, PA errors elicited stronger activation than active avoids or spurious actions, consistent with responses to the aversive US. Only Class 5 neurons showed stronger activation for active avoids than for spurious actions, indicating that this subset of neurons—strongly coupled to movement—selectively encodes the avoidance action from CS onset as this effect was also present in the orienting window. Class 5 neurons also encoded passive avoidance, showing stronger activation than during ignored Ig trials, a pattern shared to a lesser extent by Classes 1 and 3 (Fig. 9C). Comparisons between outcomes (Fig. 9D) are deemphasized because movement was not controlled across outcomes by design; accordingly, these effects likely reflect movement-related activity, which was inverted in Class 3 neurons relative to the other classes—showing stronger activation during no-action (low-movement) than during action (high-movement) trials.

In the from-action window, a significant Contingency × Outcome × Class interaction ($\chi^2(8) = 257.09$, $p < 0.0001$) demonstrated that action-related activity strongly depended on neuron class. Class 5 and 4 neurons showed stronger activation for active avoids than for spurious actions, indicating that these movement-correlated GABAergic neurons selectively encode the avoidance action itself. In both classes, passive avoid errors evoked stronger activation than active avoids, but this activation was comparable to that observed for escapes, consistent with responses driven by the aversive US. In contrast, Class 1 neurons showed the opposite relationship, with passive avoid errors producing less activation than active avoids or spurious actions. Across all classes except Class 1, escapes produced stronger activation than active avoids, further supporting a general sensitivity to the aversive US across most neuronal populations.

Together, these results indicate that during AA39, the encoding structure of PFC GABAergic neurons remains broadly like that observed in AA19 but incorporates distinct features consistent with the greater cognitive demands imposed by the added PA contingency. Baseline activity in most classes predicted subsequent errors, suggesting that elevated pre-CS excitability promotes inappropriate or mistimed responses. Class 5 neurons, which in AA19 selectively encoded the avoidance action, now showed strong activation during both correct active and passive avoidance, indicating recruitment for both forms of action control. Their activation also reflected error trials, consistent with heightened sensitivity to aversive stimulation—a feature shared by Classes 1, 2, and 4. In contrast, Class 3 neurons lost the outcome selectivity that distinguished active avoids from spurious actions in AA19, suggesting reduced contingency-specific encoding. Thus, under the more demanding conditions of AA39, Class 5 GABAergic PFC neurons—those most tightly coupled to movement—encode correct avoidance performance regardless of whether it is expressed through action or inaction.

PFC GABAergic neurons encode cautious actions

Whereas the previous analysis sorted PFC neurons based on their correlation with movement, we next asked whether PFC activity differentially encodes distinct types of active avoidance actions. To capture variation in avoidance behavior, we applied k-means clustering directly to the time series of movement speed aligned to CS onset across all correct active avoidance trials. This approach differs from the preceding one in that it classifies only correct avoidance responses based on their movement dynamics from CS onset.

This analysis revealed three distinct movement response modes (Fig. 10A, bottom gray panel). Mode 1 responses (black) were rapid onset avoids initiated immediately after the

orienting movement, which was larger than in the other modes. Mode 2 (red) and Mode 3 (cyan) responses were initiated later and exhibited significantly higher movement speed than Mode 1 when measured from-action ($p < 0.0001$), consistent with more cautious, delayed but vigorous avoidance behavior (Zhou et al., 2022).

We next used mixed-effects models to test how PFC activity varied across avoidance modes by averaging the $\Delta F/F$ time series across all recorded neurons (Fig. 10A, top panel, aligned to CS onset; Fig. 10B, aligned to avoidance action). In these models (Fig. 10C,D), the covariates retained strong effects across all windows. In the baseline and orienting windows, there were no differences across modes, indicating that the selection of avoidance strategy was not influenced by pre-CS PFC activity. During the action window, both Mode 2 ($t(4912) = 6.37$, $p < 0.0001$) and Mode 3 ($t(5246) = 4.86$, $p < 0.0001$) avoids exhibited higher activation than Mode 1 avoid, and Mode 2 tended to have higher activation than Mode 3 avoids ($t(4698) = 2.04$, $p = 0.04$). These differences were stronger in the from-action window (Mode 2 vs Mode 1: $t(4955) = 7.85$, $p < 0.0001$; Mode 3 vs Mode 1: $t(5236) = 4.27$, $p < 0.0001$; Mode 2 vs Mode 3: $t(4441) = 4.29$, $p < 0.0001$). In summary, PFC activation scales with behavioral caution consisting of delayed but fast avoidance responses (Modes 2 and 3) consistently evoking the strongest PFC activation, even after controlling for movement and baseline covariates. This indicates that PFC activity does not simply mirror motor output but instead reflects higher-order aspects of behavioral control. However, because this averaging approach treats PFC neurons as a single functional population, we next examined whether distinct GABAergic neuron activations occurred across avoidance modes.

Distinct PFC GABAergic neurons encode cue and action components

We applied k-means clustering to the ΔF/F time series of individual PFC neurons to identify distinct response profiles across avoidance modes (Fig. 11A,B). This revealed three activation types with markedly different dynamics. *Type a* activation showed minimal modulation during active avoids, whereas *Type b* and *Type c* exhibited robust yet distinct activation patterns. *Type b* activation discharged sharply at CS onset and then gradually declined, while *Type c* selectively activated around the time of the avoidance action. These distinctions were captured in the mixed-effects models. Neither activation type showed baseline differences across avoidance modes. During the orienting window, *Type b* activation at CS onset was stronger during Mode 1 avoids than during Mode 3 avoids ($t(5748) = 2.67$, $p = 0.02$). During the action window, both *Type a* and *Type c* neurons exhibited enhanced activation during the more cautious Modes 2 ($t(4871) = 3.62$, $p = 0.0008$) and 3 ($t(5169) = 3.57$, $p = 0.0008$) relative to Mode 1, whereas *Type b* showed little modulation across modes. In the from-action window, both *Type b* and *Type c* activations differentiated between avoidance modes: for *Type b*, the most cautious Mode 3 showed lower activation than Mode 2 ($t(4356) = 4.69$, $p < 0.0001$) and Mode 1 ($t(4867) = 2.83$, $p = 0.0091$); for *Type c*, Modes 2 ($t(4727) = 12.03$, $p < 0.0001$) and 3 ($t(4958) = 10.74$, $p < 0.0001$) both evoked stronger activation than Mode 1. These findings indicate that PFC GABAergic neurons exhibit complementary encoding of avoidance components: *Type b* activity encodes cue-driven alerting or preparatory processes at CS onset, whereas *Type c* activation selectively represents the execution of the avoidance action. Around the time of action, *Type b* deactivation during the most cautious Mode 3 coincides with enhanced *Type c* activation, suggesting a dynamic shift from cue-based to action-based encoding as behavioral control becomes more cautious.

PFC inactivation has little effect on cued goal-directed behavior under threat

The preceding experiments suggested that PFC activation may contribute to cued goal-directed behaviors under threat, including both active and passive avoidance. If so, optogenetic inhibition of PFC should disrupt these behaviors. To test this, we used two complementary optogenetic approaches to inhibit PFC as mice learned and performed a sequence of tasks: initial active avoidance in AA1; the transition to cautious responding in AA2, where ITCs are punished and avoidance latencies increase; and finally the more demanding AA3, in which mice must discriminate between cues to either generate or postpone an action.

First, we broadly expressed eArch3.0 in PFC neurons (n = 6) using bilateral AAV-hSyn injections and implanted optical cannulas (Fig. 1A). We avoided selectively suppressing GABAergic neurons because this would disinhibit local glutamatergic circuits and produce the opposite of the intended effect—aberrant excitation that would confound behavioral interpretation. In naïve mice, we first asked whether PFC inhibition impairs acquisition of AA1 when every trial was a CS+Light trial, using green light at three powers (10–30 mW) applied during the avoidance/escape interval. Catch trials without CS or light (NoCS) and ITCs showed typical spurious actions. Mice acquired AA1 at the same rate as controls, with no detectable impairment (Fig. 12A, black symbols). Under the same regime, mice proceeded to AA2 and showed the usual shift toward longer avoidance latencies, reflecting the emergence of caution when ITCs are punished (Fig. 12A, red). Again, no deficits were observed relative to controls.

We next trained mice in AA3 tested PFC inhibition, which includes four trial types (CS1, CS1+Light, CS2, CS2+Light) signaling active or passive responding. As in AA1 and AA2, inhibiting PFC during CS1+Light trials had little effect on the probability of active avoidance

across light powers (Fig. 12B). In contrast, passive avoidance was slightly but significantly impaired in CS2+Light trials at 10 mW ($q(18)=5.25$, $p=0.0078$) and 20 mW ($q(18)=5.67$, $p=0.0041$), with a trend at 30 mW ($q(18)=3.96$, $p=0.052$). However, the magnitude of impairment was modest (<15%). Thus, PFC inhibition produced minimal effects on cued active responding and only a small decrement in passive responding.

We then used a second approach of PFC inhibition by activating GABAergic neurons with ChR2 to generate strong and sustained cortical inhibition, confirmed by optrode recordings (Fig. 12C). We crossed Vgat-Cre with Ai32 mice and applied blue light in PFC ($n = 13$), VI ($n = 6$), or No-Opsin controls ($n = 10$). In AA1, all groups learned at the same rate when every trial was a CS+Light trial (3 mW). Performance did not differ between the three groups ($\chi^2(2) = 0.0094$, $p = 0.99$). Notably, both PFC and VI groups showed faster avoidance latencies than No-Opsin controls (PFC: $t(29) = 6.96$, $p < 0.0001$; VI: $t(31) = 4.66$ $p = 0.00011$), an effect common across the cortical regions.

In AA2, all groups showed the expected increase in avoidance latencies as caution developed when ITCs were punished. In addition, No-Opsin controls showed the typical mild performance deficit (<10%) of transitioning to AA2, but this effect was absent in PFC and VI groups, suggesting that cortical inhibition may act as a salient cue that facilitates responding, consistent with the earlier AA1 latencies.

In AA3, mice performed the four trial types after training. Light delivery during CS1 (active) trials produced little effect in any group and the latency shortening observed in AA1/2 for PFC and VI groups was no longer apparent, suggesting that it is related to the initial learning phase. However, passive responding (CS2) was significantly, albeit very weakly (<15 %), impaired in both PFC and VI groups ($p < 0.0001$), consistent with the effect also observed with

eArch. No impairment occurred in the No-Opsin controls indicating that it is the cortical inhibition driving the mild impairment.

Together, the experiments show that PFC is not required for generating cued actions under threat, even when mice must discriminate between cues and choose between action execution and deferment. Effects of PFC inhibition were small, and largely shared by VI, indicating that the PFC network is not necessary for learning or performing these goal-directed behaviors.

Discussion

Our findings reveal how PFC GABAergic neurons contribute to action selection under threat, clarifying long-standing questions about their role in cue processing, cautious responding, and outcome evaluation. Across a hierarchy of cued tasks, we found that much of the apparent task-related activity in both PFC and VI could be explained by movement or baseline fluctuations, emphasizing the importance of controlling for these confounds when interpreting cortical signals. Yet after accounting for these factors, PFC neurons—but not VI neurons—showed selective increases in activity during aversive outcomes, a signature of punished errors that may reflect the appraisal or salience of unexpected pain. Notably, simple avoidance contingencies were not encoded in PFC population activity; instead, PFC encoding emerged only under more demanding conditions in which mice had to discriminate cues and select between action and inaction. This pattern suggests that PFC GABAergic neurons become differentially engaged when behavioral demands increase and cognitive load is high, rather than when the course of action is straightforward.

Single-cell imaging further revealed distinct GABAergic neuron subpopulations, some time-locked to cue onset, and others to the action itself. Classifying neurons based on their sensitivity to movement, revealed different classes of neurons encoding avoidance primarily during the challenging situation. This encoding was broad for passive avoidance but mostly restricted to a single class for active avoidance. Despite the rich activation related to cues, action selection, and punishment under threat, broad optogenetic inhibition of PFC or VI minimally altered learning or behavioral performance, indicating that these signals do not appear necessary for executing cued actions under threat. Together, these results suggest that while PFC neurons may evaluate and contextualize threat-related actions, they are not required to generate the actions themselves highlighting the distinction between encoding and controlling behavior.

A valuable procedure for studying goal-directed contingencies

The sequence of tasks we employed provides a useful framework for probing how brain regions represent behavioral contingencies. In the initial noUS phase, mice generated spurious actions driven by baseline movement and cue salience. Once the active avoidance contingency was introduced in AA19, baseline movement and ITCs declined, and mice acquired selective, robust actions to the US-predictive CS1, with faster and more vigorous movement than spurious actions to non-predictive cues. Because CS1 can be systematically contrasted with CS2 and CS3, which remain non-predictive across these phases, and with itself in the noUS phase, the task structure provides controlled comparisons across contingencies and contexts. Introducing the passive avoidance contingency in AA39 with CS2 reduced overall action rates, including ITCs and correct actions to CS1, thereby increasing error rates relative to AA19. Action latencies also rose sharply, indicating hesitation and cautious responding as mice faced more demanding cue

discrimination and action selection under threat. This design allowed CS2 responses in AA39 to be evaluated relative to their earlier non-predictive role in AA19 and contrasted with CS1, which continued to predict the active avoidance contingency, and CS3, which remained non-predictive throughout.

Together, these behavioral findings show that mice rapidly learned distinct contingencies across progressively complex tasks and flexibly adapted their behavior as demands increased. The combination of active and passive contingencies with multiple cues and intensities created a rich, systematic framework for dissecting neural computations underlying cue processing, decision-making, cautious action selection, and movement-related influences on neural activity.

Cortical GABAergic neurons are movement-sensitive

Across all analyses, GABAergic neurons in PFC and VI were strongly modulated by ongoing movement. Cross-correlation analyses showed that both translational and rotational components of movement correlated significantly with ΔF/F, with translational speed contributing most strongly. Directional analyses further suggested that PFC GABAergic neurons do not encode head turn direction, whereas VI neurons show modest contraversive preference during ongoing movement.

In agreement with these results, every mixed-effects model used throughout this study—including those examining baseline, cue, action, and outcome trial phases—identified strong effects of movement covariates on cortical GABAergic neurons across all task conditions. Both baseline speed and window-specific speed were among the largest predictors of ΔF/F, often dwarfing the effects of the task variables themselves. This demonstrates that much of the apparent “task encoding” observed in raw traces reflects underlying movement coupling.

Together, these results indicate that cortical GABAergic neurons are intrinsically sensitive to movement, which covaries with behavioral states. As highlighted by many other studies (Musall et al., 2019; Steinmetz et al., 2019; Stringer et al., 2019; Zagha et al., 2022; Hormigo et al., 2023a), movement-related signals are sufficiently large and pervasive that careful statistical control is essential when evaluating whether neural activity genuinely reflects cue-, action-, or outcome-related processes.

PFC neurons encode punished errors independent of contingencies

A main finding across behavioral phases is that GABAergic neurons in PFC consistently increase their activity during punishment in error trials. Although VI neurons also showed US-related activation, these signals were weaker than in PFC. Thus, PFC neurons displayed a selective sensitivity to aversive outcomes after controlling for baseline neural activity and movement. Several results converge on this interpretation. First, PFC neurons showed systematic increases in $\Delta F/F$ during punished errors, including escapes in AA19 and passive-avoid errors in AA39. Second, PFC neurons responded more strongly to footshock delivered alone than to the footshock embedded with auditory white noise (full US), suggesting that unexpected and concurrently unsignaled aversive events evoke stronger neural responses than aversive events accompanied by salient signals in another modality. This pattern aligns with human studies proposing that prefrontal circuits encode the appraisal or salience of pain (Lorenz et al., 2003; Lee et al., 2024). These effects were largely absent in VI, highlighting a regional dissociation in how aversive outcomes are represented. These findings indicate that PFC GABAergic neurons exhibit selective, movement-independent encoding of punished errors, consistent with a role in evaluating aversive outcomes.

PFC encodes cued contingencies mostly in difficult situations

A second key finding is that PFC encoding of avoidance contingencies emerges only under challenging conditions, not during simple, low-error tasks. During the easy AA19 task, PFC neurons did not differentiate cues that predicted threat from those that did not, and activity during avoidance actions versus spurious actions showed little contingency-specific modulation beyond punishment-related activation. Also, the transition from noUS to AA19, where mice learn signaled active avoidance, did not reveal a selective change for the threat predictive cue, CS1, as similar changes occurred for the non-predictive cues.

In contrast, when the task became more demanding in AA39, PFC activity was different for appropriate responses, such as active and passive avoids compared to spurious actions and ignored cues, respectively. This contingency-dependent recruitment manifested in several ways. PFC neurons showed increased activity during active avoids only around the moment of action, and not during the CS, suggesting that PFC becomes engaged around the time of action. This may reflect the appearance of cautious actions in the demanding AA39 environment, which are delayed but more vigorous. However, PFC neurons also showed elevated activity during correct passive avoidance relative to correctly ignored trials, indicating recruitment during action withholding under threat. Importantly, since both action generation and deferment produced increased PFC responses in AA39, this modulation may reflect increased cognitive load rather than contingency-related encoding. Thus, population-level GABAergic PFC activity becomes engaged primarily under challenging conditions, marked by cue discrimination and hesitant action selection, consistent with a role in supporting higher cognitive load rather than directing the action itself.

Selective encoding by populations of PFC GABAergic neurons

Many studies examining PFC activity under threat, across a wide range of behavioral contexts, report neural responses to cues, actions, and outcomes (e.g. (Simon et al., 2015; Diehl et al., 2018; Maggi et al., 2018; Jercog et al., 2021; Agetsuma et al., 2023; Ho et al., 2023; Martin-Fernandez et al., 2023; Sylte et al., 2024; Gu and Johansen, 2025; Lai et al., 2025)). In our study, we directly compared neural activity for equivalent cues that differed only in contingency (e.g., threat-predictive vs. non-predictive tones varying only in frequency), and for equivalent motor responses driven either by the predictive cue or by non-predictive cues (e.g., active avoids vs. spurious crossings; passive avoids vs. correctly ignored cues). We believe these comparisons provide a more selective assessment of task-specific encoding than other common contrasts with baseline activity or general movements (e.g., ITCs lack a cue altogether). Even under these stricter conditions, a subset of neurons showed selective activation linked to correct responding, consistent with the general view that some PFC neurons encode action selection under threat.

When we classified individual PFC GABAergic neurons using ΔF/F–movement cross-correlations, we found during AA19 that only a small subset exhibited selective encoding of correct actions. One subset (Class 3), which was weakly inhibited by movement, distinguished correct actions from spurious ones when aligned to CS onset but not when aligned to action execution. A second subset (Class 5), which was strongly movement-modulated, encoded correct actions only when aligned to the action itself, not the cue. The remaining classes were broadly responsive to the aversive outcome, underscoring the strong representation of foot-shock across

the population. Thus, selective cue- and action-related encoding during simple contingencies was sparse and consistent with the predominant population-level patterns seen in photometry.

During the more challenging AA39 task, when mice learned the passive avoidance contingency, encoding patterns reorganized. Class 5 neurons continued to encode correct action generation and additionally represented correct action deferment. Class 1 and Class 3 also represented the correct passive responding compared to correctly ignored cues. Encoding of aversive outcomes remained widespread across classes, paralleling the photometry results.

When we clustered the movement profile of correct avoidance actions, different modes or styles were revealed including cautious responding where the PFC activation was stronger. Further classifying the neural activity for each avoidance mode revealed two types of activations related to the cue and related to the action execution. The latter became more prominent during cautious responding, characterized by delayed but vigorous actions typical of high task difficulty.

Across task contingencies and increasing behavioral demands, PFC GABAergic neurons integrated contextual, sensory, and motor information. Their selective engagement increased under high cognitive load and uncertainty, while broad activity reflected evaluation of aversive outcomes and the processing of cautious responding under threat.

Is PFC required for cued actions under threat?

Although PFC involvement is frequently cited in cued actions under threat, including what is broadly termed "*avoidance*," the term itself is used inconsistently in neuroscience and often refers to procedures that measure distinct constructs. For example, some studies use conditioned place preference/aversion, which assesses how animals evaluate contexts associated with reward or aversion rather than whether they can generate or withhold actions to avoid harm.

Other studies use operant lever-press avoidance tasks in rodents, which are not species-specific defensive reactions like running away and have been shown to be essentially unlearnable (Bolles, 1970). Still others investigate conflict between approach and avoidance using a platform task (Diehl et al., 2018). In this context, conflict refers to the competition between opposing motivational demands, such as crossing a dangerous, busy road to reach a water fountain while thirsty, which is distinct from non-conflictual avoidance or approach (Miller, 1944; Botvinick et al., 2004; Aupperle et al., 2015). Finally, considering the species is also important because although mice and rats share threat-related behaviors they also exhibit differences (Blanchard et al., 2001, 2003, 2005).

Focusing specifically on classical active avoidance in rodents (Mowrer, 1939), lesion findings have been mixed. In rats, pre-training PFC lesions slow learning whereas post-training lesions have no effect (Castro-Alamancos and Borrell, 1992). Another rat study reported similar slowing only when lesions included both dorsomedial and infralimbic PFC (not dorsomedial alone), while post-training lesions were not examined (Moscarello and LeDoux, 2013). In mice, some studies report that dorsomedial PFC inactivation impairs performance (Jercog et al., 2021), whereas others find minimal effects of irreversible lesions across avoidance contingencies (Zhou et al., 2022) in agreement with the optogenetic inactivation in the present study. In conflict procedures, there have also been inconsistencies even within the same study, such as muscimol inactivation delaying avoidance, while optogenetic inhibition is ineffective (Diehl et al., 2018). In our study, action generation and deferment were never in direct conflict. We found that when baseline and movement variables are controlled in statistical models, encoding of avoidance contingencies is limited. Consistent with this sparse encoding, optogenetic inactivation of PFC had minimal effects on avoidance in a variety of tasks of increasing difficulty.

Materials and Methods

Experimental Design and Statistical Analysis

The methods used in the present paper were like those employed in our previous studies (e.g., (Hormigo et al., 2023b; Zhou et al., 2024, 2025)). All procedures were reviewed and approved by the institutional animal care and use committee and conducted in adult (>8 weeks) male and female mice. Most experiments used a repeated-measures design in which each mouse or cell served as its own control (within-subject/cell comparisons), but we also compared experimental groups between subjects or cells (between-group comparisons). To test the main effects of experimental variables, we used either repeated-measures ANOVA or linear mixed-effects models.

Mixed-effects models were used to quantify relationships between ΔF/F activity (dependent variable) and experimental factors while accounting for repeated measures. Models included fixed effects for task contingency, outcome, cell/avoid classes, etc., and their interactions, along with window-specific covariates for head speed and baseline ΔF/F (centered using z-scores as noted in the results). Random effects were specified as sessions nested within subjects or cells nested within subjects, following the general syntax: ΔF/F ~ (Factor1 * Factor2 * …) * Speed + (1 | Subject/Session) in lme4. Separate models were fit for each of the four behavioral windows, using area measures for the baseline, orienting, avoidance, and from-action windows (normalized by window duration), with the latter windows corrected by the baseline. Generally, the baseline window (-0.5 to 0 s from CS onset) captures the pre-cue state at trial initiation, whereas the orienting window (0 to 0.5 s post-CS) reflects the initial response to the cue. The avoidance window (with task-specific durations) corresponds to the period when mice must generate or withhold a response to avoid punishment, excluding both the preceding

orienting window and the subsequent escape interval. The from-action window aligns time series from the occurrence of a response (e.g., avoid, escape, passive avoid error), enabling comparison between avoidance responses occurring within the avoidance interval and escape responses in the later interval. Including the window-specific head speed covariates, along with baseline $\Delta F/F$ and head speed measures, in the post-CS window models allowed us to dissociate movement-related modulation from neural encoding of task-related factors. Model fit was assessed by likelihood ratio tests comparing nested models, confirming that inclusion of covariates and interactions improved explanatory power. For significant main effects or interactions, post-hoc pairwise comparisons were performed using *emmeans* with Tukey's correction (Holm-adjusted). For continuous covariates, slope effects were evaluated using *emtrends*, with significance determined by t-tests ($p < 0.05$). Equivalence of movement measures was assessed using two one-sided tests (TOST) with a tolerance (delta) of ±6 cm/s (<50 % of baseline speed and <10% of peak speed) to confirm that classified neuronal activity differences were not generally attributable to movement differences.

To enable rigorous approaches, we maintain a centralized metadata system that logs all details about the experiments and is engaged for data analyses (Castro-Alamancos, 2022). Moreover, during daily behavioral sessions, computers run experiments automatically using preset parameters logged for reference during analysis. Analyses are performed using scripts that automate all aspects of data analysis from access to logged metadata and data files to population statistics and graph generation.

Strains and Adeno-Associated Viruses (AAVs)

To record from GABAergic PFC or VI neurons using calcium imaging, we injected a Cre-dependent AAV (AAV5-syn-FLEX-jGCaMP7f-WPRE (Addgene: $7x10^{12}$ vg/ml) in the PFC or VI of Vgat-cre mice (Jax 028862; B6J.129S6(FVB)-Slc32a1$^{tm2(cre)Lowl}$/MwarJ) to express GCaMP6f/7f.

To excite GABAergic neurons in PFC or VI using optogenetics, we crossed Vgat-cre and Ai32 (Jax 024109; B6.Cg-Gt(ROSA)26Sor$^{tm32(CAG-COP4*H134R/EYFP)Hze}$/J) mice (Vgat-Chr2 mice). To broadly inhibit PFC neurons using optogenetics, we expressed eArch3.0 by injecting AAV5-hSyn-eArch3.0-EYFP (UNC Vector Core, titers: $5.9x10^{12}$ vg/ml) in the PFC of C57BL/6 mice (PFC-Arch mice). The optogenetic methods used in the present study have been validated in previous studies using slice and/or in vivo electrophysiology (Hormigo et al., 2016, 2019, 2021). No-Opsin controls were injected with AAV8-hSyn-EGFP (Addgene, titers: $4.3x10^{12}$ GC/ml by quantitative PCR) or nil in the PFC.

Surgeries

Optogenetics and fiber photometry surgeries were performed under isoflurane anesthesia (~1%), with postoperative carprofen administered for analgesia. AAVs (0.4 µl per site) were injected into the PFC using stereotaxic coordinates given in millimeters relative to bregma (anteroposterior [AP], mediolateral [ML], dorsoventral [DV] below the bregma-lambda plane). For both Arch and GCaMP experiments, injections were made at AP = 2.0, ML = ±0.3, and DV = 1.5-2. Injections were unilateral for GCaMP and bilateral for Arch. Optical fibers were implanted at the time of surgery in the location of the AAV injections or centered on the midline: for fiber photometry, a 400 µm diameter fiber was placed unilaterally at ML = 0.3 mm; for

miniscope recordings, a 600 µm diameter GRIN lens was placed unilaterally at ML = 0.3 mm; and for optogenetic experiments, a dual 200 µm diameter fiber was employed.

For optogenetics or fiber photometry experiments targeting VI, dual optogenetic cannulas or single fiber photometry implants were placed at AP = –2.6, ML = ±2.0, and DV = 0.2, with GCaMP AAV injections at the same coordinate as the cannula. Control animals without opsins ("No Opsin" mice) were implanted with cannulas in the PFC or adjacent areas, and their data were pooled after confirming that light delivery produced similar behavioral effects across animals.

Active Avoidance tasks

Mice are placed in a standard shuttle box (16.1" x 6.5") that has two compartments separated by a partition with side walls forming a doorway that the animal must traverse to shuttle between compartments. A speaker is placed on one side, but the sound fills the whole box and there is no difference in behavioral performance (signal detection and response) between sides. AA19 and AA39 sessions (n=7 each) are updated versions of previously described tasks (AA1 and AA3; e.g., (Zhou et al., 2022, 2023)) using three cues (CS1, CS2, CS3; 8, 4, 12 kHz tones) applied at three intensities (65-87 dB) to manipulate saliency and increase task difficulty. Thus, AA1 and AA3 are the same tasks as AA19 and AA39 using one or two salient cues respectively, while AA2 is like AA1 but punishes ITCs (0.2 s US).

During noUS sessions, the three CS were presented randomly without consequence during 7 sec windows, but crossing compartments (action) turned off the CS. During AA19, mice were presented with the same cues, but CS1 incorporated the active avoidance contingency. During AA39, the sessions continue as during AA19, but CS2 includes the passive avoidance

contingency. In AA19/39, mice are free to cross between compartments during the intertrial interval; there is no consequence for intertrial crossings (ITCs).

CS1 trials in AA19/39 consist of a 7 sec avoidance interval followed by a 10 sec escape interval. During the avoidance interval, an auditory CS (8 kHz) is presented for the duration of the interval or until the animal produces a conditioned response (avoidance response) by moving to the adjacent compartment, whichever occurs first. If the animal avoids, by moving to the next compartment, the CS ends, the escape interval is not presented, and the trial terminates. However, if the animal does not avoid, the escape interval ensues by presenting white noise and a mild scrambled electric foot-shock (0.3 mA) delivered through the grid floor of the occupied half of the shuttle box. This unconditioned stimulus (US) readily drives the animal to move to the adjacent compartment (escape response), at which point the US terminates, and the escape interval and the trial ends. Thus, an *avoidance response* will eliminate the imminent presentation of a harmful stimulus. An *escape response* is driven by presentation of the harmful stimulus to eliminate the harm it causes. Successful avoidance warrants the absence of harm. Each trial is followed by an intertrial interval (duration is randomly distributed; 25-45 sec range), during which the animal awaits the next trial.

In AA39, mice continue to be randomly presented the three cues but are subjected to a CS discrimination procedure in which they must respond differently to a CS1 and a CS2. Mice perform the basic signaled active avoidance to CS1 but also perform signaled passive avoidance to CS2. In AA39, if mice shuttle during the CS2 avoidance interval (7 sec), they receive a 0.5 sec foot-shock (0.3 mA) with white noise and the trial ends. If animals do not shuttle during the CS2 avoidance interval, the CS2 trial terminates at the end of the avoidance interval (i.e., successful

signaled passive avoidance). AA1/2 and AA3 are the same tasks as AA19 and 39 using salient cues.

There are three main variables representing task performance. The percentage of active avoidance responses (% avoids) represents the trials in which the animal actively avoided the US in response to the CS. The response latency (latency) represents the time (sec) at which the animal enters the safe compartment after the CS onset; avoidance latency is the response latency only for successful active avoidance trials (excluding escape trials). The number of crossings during the intertrial interval (ITCs) represents random shuttling due to locomotor activity in the AA19 and AA39 procedures. The sound pressure level (SPL) of the auditory CS's were measured using a microphone (PCB Piezotronics 377C01) and amplifier (x100) connected to a custom LabVIEW application that samples the stimulus within the shuttle cage as the microphone rotates driven by an actuator controlled by the application.

Fiber photometry

We employed a 2-channel excitation (465 and 405 nm) and 2-channel emission (525 and 430 nm for GCaMP6f and other emissions) fiber photometry system (Doric Lenses). Alternating light pulses were delivered at 100 Hz (per each 10 ms, 465 is on for 3 ms, and 2 ms later 405 is on for 3 ms). While monitoring the 525 nm emission channel, we set the 465 light pulses in the 20-60 µW power range and then the power of the 405 light pulses was adjusted (20-50 µW) to approximately match the response evoked by the 465 pulses. During recordings, the emission peak signals evoked by the 465 (GCaMP6f) and 405 (isobestic) light pulses were acquired at 5-20 kHz and measured at the end of each pulse. To calculate F_o, the measured peak emissions evoked by the 405 nm pulses were scaled to those evoked by the 465 pulses (F) using the slope

of the linear fit. Finally, F/F_o was calculated with the following formula: $(F-F_o)/F_o$ and converted to Z-scores. Due to the nature of the behavior studied, a swivel is essential. We employed a rotatory-assisted photometry system that has no light path interruptions (Doric Lenses). In addition, charcoal powder was mixed in the dental cement to assure that ambient light was not leaking into the implant and reaching the optical fiber; this was tested in each animal by comparing fluorescence signals in the dark versus normal cage illumination.

Miniscope imaging

We employed GRIN lenses (0.6mm diameter, 7.3mm length) and an nVista (Inscopix) recording system coupled with an electrical swivel. We added custom tungsten rods to the GRIN lenses to maximize the stability of the recordings. During each session, we adjusted the focus plane to record the best neurons, which were assigned as neurons per session. We used Inscopix Data Processing Software (IDPS) to extract ROIs from the raw miniscope movies. Briefly, the movies were preprocessed with a spatial bandpass filter, which removes the low and high spatial frequency content from the movies minimizing out of plane neuropil fluorescence and allows visual identification of putative neurons. The movies were then motion corrected, using the initial frame as the global reference. Finally, neurons were manually outlined on the processed movie, and the F/F calcium activity was calculated from these ROIs. To assure stability during each recording session, the ROIs were visually verified from start to end for each video.

K-means clustering (scikit-learn) was performed on features extracted from F/F (z-score) or movement (speed) time series traces. These features included peak amplitudes, areas under the curve, and times to peak around specific events (e.g., turns, CS onset, or avoidance occurrence).

Alternatively, clustering was also applied to principal component scores derived from the same traces, which were treated as an analog of spectral data for PCA.

Optogenetics

The implanted optical fibers were connected to patch cables using sleeves. A black aluminum cap covered the head implant and completely blocked any light exiting at the ferrule's junction. Furthermore, the experiments occurred in a brightly lit cage that made it difficult to detect any light escaping the implant. The other end of the patch cable was connected to a dual light swivel (Doric lenses) that was coupled to a green laser (520 nm; 100 mW) to activate Arch or a blue laser (450 nm; 80 mW) to activate ChR2. In experiments expressing Arch, Green light was applied continuously at different powers (10, 20, and 30 mW) during the indicated intervals. In experiments with ChR2 expression, blue light was delivered continuously at constant power (2-3 mW). Power is regularly measured by flashing the connecting patch cords onto a light sensor—with the sleeve on the ferrule.

During optogenetic experiments that involve avoidance procedures, we compared different trial types: CS and CS+Light. *CS trials* were standard avoidance trials specific to each procedure, without optogenetic stimulation. *CS+Light trials* were identical to CS trials, except that optogenetic light was delivered simultaneously with the CS and the US during the avoid and escape intervals. Catch trials present a blank no-CS trial without consequence to gauge random responding, which is also evaluated with ITCs. To perform within group repeated measures (RM) comparisons, the different trial types for a procedure were delivered randomly within the same session. In addition, the trials were compared between different groups, including No

Opsin mice that did not express opsins but were subjected to the same trials including light delivery.

Video tracking

All mice in the study (open field or shuttle box) were continuously video tracked (30-100 FPS) in synchrony with the procedures and other measures. During open field experiments, mice are placed in a circular open field (10" diameter) that was illuminated from the bottom or in the standard shuttle box (16.1" x 6.5"). We automatically tracked head movements with two color markers attached to the head connector –one located over the nose and the other between the ears. The coordinates from these markers form a line (head midline) that serves to derive several instantaneous movement measures per frame (Zhou et al., 2023). Overall head movement was separated into *rotational* and *translational* components (unless otherwise indicated, overall head movement is presented for simplicity and brevity, but the different components were analyzed). Rotational movement was the angle formed by the head midline between succeeding video frames multiplied by the radius. Translational movement resulted from the sum of linear (forward vs backward) and sideways movements. *Linear* movement was the distance moved by the ears marker between succeeding frames multiplied by the cosine of the angle formed by the line between these succeeding ear points and the head midline. *Sideways* movement was calculated as linear movement, but the sine was used instead of the cosine. Pixel measures were converted to metric units using calibration and expressed as speed (cm/sec). We used the time series to extract window measurements around events (e.g., CS presentations). Measurements were obtained from single trial traces and/or from traces averaged over a session. In addition, we obtained the direction of the rotational movement with a *Head Angle* or *bias* measure, which was

the accumulated change in angle of the head per frame (versus the previous frame) zeroed by the frame preceding the stimulus onset or event (this is equivalent to the rotational speed movement in degrees). The *time to peak* is when the *extrema* occurs versus event onset.

To detect spontaneous turns or movements from the head tracking, we applied a local maximum algorithm to the continuous head angle or speed measure, respectively. Every point is checked to determine if it is the maximum or minimum among the points in a range of 0.5 sec before and after the point. A change in angle of this point >10 degrees was a detected turn in the direction of the sign. We further sorted detected turns or movements based on the timing of previous detected events.

Histology

Mice were deeply anesthetized with an overdose of isoflurane. Upon losing all responsiveness to a strong tail pinch, the animals were decapitated, and the brains were rapidly extracted and placed in fixative. The brains were sectioned (100 µm sections) in the coronal or sagittal planes. All sections were mounted on slides, cover slipped with DAPI mounting media, and all the sections were imaged using a slide scanner (Leica Thunder). We used an APP we developed with OriginLab (Brain Atlas Analyzer) to align the sections with the Allen Brain Atlas Common Coordinate Framework (CCF) v3 (Wang et al., 2020). This reveals the location of probes and fluorophores versus the delimited atlas areas.

Figure Legends

Figure 1. Calcium imaging fiber photometry reveals that both PFC and VI GABAergic neurons activate during spontaneous exploratory movement.

A, Parasagittal sections showing the optical fiber tract reaching PFC and VI with GCaMP7f fluorescence expressed in GABAergic neurons around the fiber ending. The section was aligned with the Allen brain atlas CCF. The right panels show schematics of the locations of optical fiber endings in PFC and VI.

B, ΔF/F calcium imaging time extracted around detected spontaneous movements recorded in PFC (left) or VI (right). Time zero represents the peak of the movement. The upper traces show ΔF/F mean±SEM of all movement peaks (black), those that had no detected peaks 3 s prior (red), and peaks taken at a fixed interval >5 s (cyan). The lower traces show the corresponding movement speed for the selected peaks. All traces in the paper are mean±SEM.

C, Cross-correlation between movement and ΔF/F for the overall (left), translational (middle) and rotational (right) components in PFC (red) and VI (cyan).

D, Per session (dots) and mean±SEM (rectangle) linear fit (correlation, r) between overall movement and ΔF/F in PFC (left) and VI (right), including the rotational and translational components. The lighter dots show the linear fit after scrambling one of the variables (shuffled).

Figure 2. PFC and VI GABAergic neurons differentially encode exploratory movement turning direction.

A, ΔF/F calcium imaging, overall movement, rotational movement, and angle of turning direction for detected movements classified by the turning direction (ipsiversive and contraversive; red and cyan) versus the side of the recording (implanted optical fiber) in PFC. At time zero, the animals spontaneously turn their head in the indicated direction. The columns show all turns (left), those that included no turn peaks 3 s prior (middle), and peaks selected at a fixed interval >5 s (right). Note that the speed of the movements was similar in both directions (the y-axis speed is truncated to show the rising phase of the movement).

B, Same as A for VI.

C, Population measures (area, peak amplitude, and time to peak for traces 3 s around the detected peaks) of ΔF/F and overall movement for the different classified peaks in PFC. Asterisks denote significant differences ($p<0.05$) between ipsiversive and contraversive movements.
D, Same as C for VI.

Figure 3. Behavioral performance in the cue-driven action tasks employed in the study.
A, Arrangement of the shuttle box used for the tasks.
B, Behavioral performance during the three procedures that occur sequentially (noUS, AA19, and AA39) for 5-7 sessions each. The panels show the percentage of actions (shuttle box crossings) (top), action latency (middle), and intertrial crossings (ITCs) (lower). Three different auditory cues (CS1-3) are presented randomly, each at three intensities (low, medium, high dB). During noUS, the CSs signal no aversive consequence, and crossings are spurious actions that turn off the CS, while not acting are ignored cues. During AA19, CS1 signals that the mice must act to avoid an aversive US, while CS2 and CS3 remain non-predictive. If mice do not avoid during CS1, the US drives a rapid escape. During AA39, CS2 signals the need to withhold action, passively avoiding a brief aversive US.
C, Same data in B averaging across CS intensities.

Figure 4. Movement traces (overall speed) across task phases (noUS, AA19, AA39), cues (CS1–CS3), and cue intensities (low, medium, high), separated by action (***A***) and no-action (***B***) outcomes for all PFC and VI mice. Dashed boxes indicate correct or error responses for the active avoidance (AA) and passive avoidance (PA) contingencies; traces outside the boxes correspond to spurious actions or correctly ignored cues for the Ig contingency. Note that escapes are AA errors. Each panel overlays the three cue intensities. Some traces are intentionally truncated to highlight specific features, but the full trace is shown in C.
C, Movement for the action trials in A, replotted aligned to the action occurrence ("from-action").

Figure 5. ΔF/F fiber photometry traces and corresponding movement traces for PFC mice across task phases (noUS, AA19, AA39) and cues (CS1–CS3), separated by action (***A***) and no-action (***B***) outcomes. For each cue (CS1, CS2, CS3; left, middle, right panels), traces from the three

task phases are overlaid. Movement traces for each condition are shown in the lower panels. Arrows indicate correct or error responses for the active-avoidance (AA) and passive-avoidance (PA) contingencies. Some traces are intentionally truncated to highlight specific features, but the full trace is shown in C.

C, Traces for the action trials in A, replotted aligned to the action occurrence ("from-action").

Figure 6. Population estimates of marginal means from mixed-effects models. Movement-only models across all mice are shown in ***A-B***, and ΔF/F fiber-photometry models (including covariates) are shown for PFC (***C-D***) and VI (***E-F***) neurons. Left panels (***A, C, E***) compare trial types across task phases; right panels (***B, D, F***) compare action and no-action outcomes within each task and CS. In the left panels, asterisks indicate significant differences relative to the preceding task phase. In the right panels, asterisks indicate significant differences between action and no-action outcomes. Each panel set contains four rows corresponding to the baseline, orienting, action, and from-action model windows. Transparency in the action window indicates that movement was not controlled (held constant) between outcomes in that model.

Figure 7. Population estimates of marginal means from simplified ΔF/F fiber-photometry mixed-effects models (including covariates) for AA19 (***A-B***) and AA39 (***C-D***) tasks models, shown separately for PFC neurons (**A, C**) and VI neurons (***B, D***). Each panel set contains four rows corresponding to the baseline, orienting, action, and from-action model windows. Within each set, the left and middle panels compare contingencies (Ignore, Active Avoid, Passive Avoid; Ig, AA, PA) for action and no-action trials. The right panel compares outcomes (action vs. no-action). Transparency in the Action window indicates that movement was not controlled (held constant) between outcomes in that model.

E, ΔF/F fiber-photometry (top) and movement (bottom) traces for the unsignaled escape task in PFC and VI neurons, showing overlaid US, foot-shock, and white noise (WN) trials aligned to the US onset.

F, Population estimates of marginal means from the ΔF/F fiber-photometry mixed-effects model (including movement covariates) for the action window.

Figure 8. Different classes of PFC GABAergic neurons classified by cross-correlation between speed and ΔF/F during the AA19 task.

A, ΔF/F fiber-photometry traces (top) and movement traces (bottom) for five neuron classes identified by applying k-means clustering to the cross-correlation functions (inset **a**). Class 1 (black) and Class 2 (red) neurons showed moderate positive cross-correlation with movement. Class 3 (green) showed moderate inhibition around movement. Class 4 (blue) and Class 5 (cyan) exhibited strong positive cross-correlations, with Class 5 activity preceding Class 4 around zero lag. Each panel shows traces for the three CSs (CS1–CS3), split by outcome (action vs no-action), which may be correct, errors, or spurious depending on the contingency. In AA19, CS1 signals the need to actively avoid.

B, Same as A, but aligned from-action. There are no traces from-action for no-action trials.

C, Marginal mean ΔF/F estimates from linear mixed-effects models for the baseline, orienting, avoidance, and from-action windows, comparing the Ignore (Ig) and Active Avoid (AA) contingencies across outcomes (action vs. no-action) for the five neuron classes.

D, Same as C, but comparing outcomes within each contingency for each class. Transparency in the action window indicates that movement was not controlled (held constant) between outcomes in that model.

Figure 9. Different classes of PFC GABAergic neurons classified by cross-correlation between speed and ΔF/F during the AA39 task.

A, ΔF/F fiber-photometry traces (top) and movement traces (bottom) for five neuron classes identified by applying k-means clustering to the cross-correlation functions (inset **a**). Class 1 (black) and Class 2 (red) neurons showed moderate positive cross-correlation with movement. Class 3 (green) showed moderate inhibition around movement. Class 4 (blue) and Class 5 (cyan) exhibited strong positive cross-correlations, with Class 5 activity preceding Class 4 around zero lag. Each panel shows traces for the three CSs (CS1–CS3), split by outcome (action vs no-action), which may be correct, errors, or spurious depending on the contingency. In AA39, CS1 signals the need to actively avoid, whereas CS signals the need to passively avoid..

B, Same as A, but aligned from-action. There are no traces from-action for no-action trials.

C, Marginal mean ΔF/F estimates from linear mixed-effects models for the baseline, orienting, avoidance, and from-action windows, comparing the Ignore (Ig), Active Avoid (AA), and

passive avoid (PA) contingencies across outcomes (action vs. no-action) for the five neuron classes.

D, Same as C, but comparing outcomes within each contingency for each class. Transparency in the action window indicates that movement was not controlled (held constant) between outcomes in that model.

Figure 10. Activation of PFC GABAergic neurons across three distinct signaled active avoidance modes.

A, k-means clustering of movement speed time-series from CS onset (gray panel) revealed three distinct active avoidance modes. Mode 1 avoids were initiated rapidly after CS onset, whereas Mode 2 and Mode 3 avoids were delayed, reflecting increasingly cautious responding. The top panel shows the average ΔF/F activity of all recorded PFC neurons for each avoidance mode. PFC activation was weaker during Mode 1, compared to Mode 2 and Mode 3.

B, Same as in A, but aligned from-action (avoid) occurrence.

C, Marginal means (ΔF/F) from the linear mixed-effect models for the baseline, orienting, avoidance and from-action windows for the three active avoidance Modes.

Figure 11. Three distinct subtypes of PFC GABAergic neuron activation signaled active avoidance mode representing the CS or the avoidance action for equivalent movements.

A, k-means clustering of the ΔF/F time series within each avoid mode (from Fig. 10) revealed three activation subtypes (a-c). Type a (1-3a) showed little activation across modes. Type b (1-3b) activated sharply starting from CS onset. Type c (1-3c) displayed selective activation in relation to the avoidance movement. Type b activation was strongest for Mode 1 and demodulated in Mode 3, while Type c activation was more robust in Modes 2 and 3. Traces are aligned from CS onset.

B, Same data as in A, aligned from-action.

C, Marginal means (ΔF/F) from the linear mixed-effect models for the baseline, orienting, avoidance and from-action windows for the active avoid modes and activation types.

Figure 12. Optogenetic inhibition of PFC neurons has little effect on signaled avoidance.

A, Effects of continuous green light at three different powers (10-30 mW) on learning AA1 (black circles) followed by AA2 (red circles) in naïve mice expressing eArch3.0 in PFC neurons. In AA1, the CS signals the active avoidance contingency. In AA2, ITCs are punished. On every trial, light was delivered during the CS and US to inhibit PFC activity. Mice learned and performed the task like control mice, which are not shown (but see D, E). During AA2, performance remained unchanged, although mice exhibited the typical shift toward longer avoidance latencies. Catch trials (NoCs) consisted of blank trials without light or consequence for crossing.

B, Same as A, but during subsequent AA3 task, in which CS1, CS1+Light, CS2, and CS2+Light trials were presented randomly. Light trials included the three powers used in A. CS1 signals the active avoidance (AA) contingency, whereas CS2 signals the passive avoidance (PA) contingency.

C, Example PSTH showing the effect of blue light in Vgat-Cre mice expressing ChR2 in GABAergic neurons recorded in freely behaving mice with an optrode implanted in PFC. Continuous blue light for 1 s produces virtually complete sustained abolishment of multi-unit activity (MUA). The histogram is the average of 15 light trials. The same effects were observed in other mice (n=3).

D, Effects of continuous blue light in PFC on learning AA1 followed by AA2 in naïve mice expressing ChR2 in GABAergic neurons. The columns represent groups of mice with optical cannulas in PFC or VI, with a third No Opsin group with optical fibers in PFC that did not express ChR2. The AA1 and AA2 tasks are like those in A, including catch trials. The right panel shows the contrasts between AA1 and AA2 for each group.

E, Effect of blue light in the same three groups during AA3.

References

Agetsuma M, Sato I, Tanaka YR, Carrillo-Reid L, Kasai A, Noritake A, Arai Y, Yoshitomo M, Inagaki T, Yukawa H, Hashimoto H, Nabekura J, Nagai T (2023) Activity-dependent organization of prefrontal hub-networks for associative learning and signal transformation. Nat Commun 14:5996.

Aupperle RL, Melrose AJ, Francisco A, Paulus MP, Stein MB (2015) Neural substrates of approach-avoidance conflict decision-making. Hum Brain Mapp 36:449–462.

Barack DL, Miller EK, Moore CI, Packer AM, Pessoa L, Ross LN, Rust NC (2022) A call for more clarity around causality in neuroscience. Trends Neurosci 45:654–655.

Blanchard DC, Blanchard RJ, Griebel G (2005) Defensive responses to predator threat in the rat and mouse. Curr Protoc Neurosci Chapter 8:Unit 8 19.

Blanchard DC, Griebel G, Blanchard RJ (2001) Mouse defensive behaviors: pharmacological and behavioral assays for anxiety and panic. Neurosci Biobehav Rev 25:205–218.

Blanchard DC, Griebel G, Blanchard RJ (2003) The Mouse Defense Test Battery: pharmacological and behavioral assays for anxiety and panic. Eur J Pharmacol 463:97–116.

Bogacz R, Wagenmakers EJ, Forstmann BU, Nieuwenhuis S (2010) The neural basis of the speed-accuracy tradeoff. Trends Neurosci 33:10–16.

Bolles RC (1970) Species-specific defense reactions and avoidance learning. Psychological Review 77:32–48.

Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. Trends in Cognitive Sciences 8:539–546.

Castro-Alamancos MA (2004a) Dynamics of sensory thalamocortical synaptic networks during information processing states. Prog Neurobiol 74:213–247.

Castro-Alamancos MA (2004b) Absence of rapid sensory adaptation in neocortex during information processing states. Neuron 41:455–464.

Castro-Alamancos MA (2022) A System to Easily Manage Metadata in Biomedical Research Labs Based on Open-source Software. Bio Protoc 12:e4404.

Castro-Alamancos MA, Borrell J (1992) Facilitation and recovery of shuttle box avoidance behavior after frontal cortex lesions is induced by a contingent electrical stimulation in the ventral tegmental nucleus. Behavioural Brain Research 50:69–76.

Castro-Alamancos MA, Connors BW (1996) Short-term plasticity of a thalamocortical pathway dynamically modulated by behavioral state. Science 272:274–277.

Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499:295–300.

Diehl MM, Bravo-Rivera C, Rodriguez-Romaguera J, Pagan-Rivera PA, Burgos-Robles A, Roman-Ortiz C, Quirk GJ (2018) Active avoidance requires inhibitory signaling in the rodent prelimbic prefrontal cortex. eLife 7:e34657.

Euston DR, Gruber AJ, McNaughton BL (2012) The role of medial prefrontal cortex in memory and decision making. Neuron 76:1057–1070.

Friedman NP, Robbins TW (2022) The role of prefrontal cortex in cognitive control and executive function. Neuropsychopharmacology 47:72–89.

Gu X, Johansen JP (2025) Prefrontal encoding of an internal model for emotional inference. Nature 643:1044–1056.

Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ (2012) Go and no-go learning in reward and punishment: Interactions between affect and effect. NeuroImage 62:154–166.

Gupta A, Wang Y, Markram H (2000) Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex [see comments]. Science 287:273–278.

Heitz RP, Schall JD (2012) Neural mechanisms of speed-accuracy tradeoff. Neuron 76:616–628.

Ho Y-Y, Yang Q, Boddu P, Bulkin DA, Warden MR (2023) Ventromedial prefrontal parvalbumin neurons are necessary for initiating cued threat avoidance. eLife 12.

Hormigo S, Shanmugasundaram B, Zhou J, Castro-Alamancos MA (2021) A Signaled Locomotor Avoidance Action Is Fully Represented in the Neural Activity of the Midbrain Tegmentum. J Neurosci 41:4262–4275.

Hormigo S, Vega-Flores G, Castro-Alamancos MA (2016) Basal Ganglia Output Controls Active Avoidance Behavior. J Neurosci 36:10274–10284.

Hormigo S, Vega-Flores G, Rovira V, Castro-Alamancos MA (2019) Circuits That Mediate Expression of Signaled Active Avoidance Converge in the Pedunculopontine Tegmentum. J Neurosci 39:4576–4594.

Hormigo S, Zhou J, Chabbert D, Sajid S, Busel N, Castro-Alamancos M (2023a) Zona incerta distributes a broad movement signal that modulates behavior. Elife 12:RP89366.

Hormigo S, Zhou J, Chabbert D, Sajid S, Busel N, Castro-Alamancos M (2023b) Zona incerta distributes a broad movement signal that modulates behavior. Elife 12:RP89366.

Jercog D, Winke N, Sung K, Fernandez MM, Francioni C, Rajot D, Courtin J, Chaudun F, Jercog PE, Valerio S, Herry C (2021) Dynamical prefrontal population coding during defensive behaviours. Nature 595:690–694.

Jonas E, Kording KP (2017) Could a Neuroscientist Understand a Microprocessor? PLOS Computational Biology 13:e1005268.

Kawaguchi Y, Kubota Y (1997) GABAergic cell subtypes and their synaptic connections in rat frontal cortex. Cerebral Cortex 7:476–486.

Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017) Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 93:480–490.

Lai C-H, Park G, Xu P, Sun X, Ge Q, Jin Z, Betts S, Liu X, Liu Q, Simha R, Zeng C, Lu H, Du J (2025) Decoding the hidden variabilities in mPFC descending pathways across emotional states. eLife 14.

Le Merre P, Ährlund-Richter S, Carlén M (2021) The mouse prefrontal cortex: Unity in diversity. Neuron 109:1925–1944.

Lee SA, Lee J-J, Han J, Choi M, Wager TD, Woo C-W (2024) Brain representations of affective valence and intensity in sustained pleasure and pain. Proceedings of the National Academy of Sciences 121:e2310433121.

Lorenz J, Minoshima S, Casey KL (2003) Keeping pain out of mind: the role of the dorsolateral prefrontal cortex in pain modulation. Brain 126:1079–1091.

Maggi S, Peyrache A, Humphries MD (2018) An ensemble code in medial prefrontal cortex links prior events to outcomes during learning. Nat Commun 9:2204.

Mark L, Linda MA, Kyra S, Samantha RW (2018) What, If Anything, Is Rodent Prefrontal Cortex? eneuro 5:ENEURO.0315-18.2018.

Martin-Fernandez M, Menegolla AP, Lopez-Fernandez G, Winke N, Jercog D, Kim H-R, Girard D, Dejean C, Herry C (2023) Prefrontal circuits encode both general danger and specific threat representations. Nat Neurosci 26:2147–2157.

Miller NE (1944) Experimental studies of conflict. In: Personality and the behavioural disorders (Huny JM, ed), pp 431–465. New York: Ronald Press.

Moscarello JM, LeDoux JE (2013) Active avoidance learning requires prefrontal suppression of amygdala-mediated defensive reactions. J Neurosci 33:3815–3823.

Mowrer OH (1939) A stimulus-response analysis of anxiety and its role as a reinforcing agent. Psychological Review 46:553–565.

Musall S, Kaufman MT, Juavinett AL, Gluf S, Churchland AK (2019) Single-trial neural dynamics are dominated by richly varied movements. Nat Neurosci 22:1677–1686.

Pisanello M, Pisano F, Hyun M, Maglie E, Balena A, De Vittorio M, Sabatini BL, Pisanello F (2019) The Three-Dimensional Signal Collection Field for Fiber Photometry in Brain Tissue. Front Neurosci 13:82.

Ratcliff R, Frank MJ (2012) Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. Neural computation 24:1186–1229.

Ridderinkhof KR, van den Wildenberg WPM, Segalowitz SJ, Carter CS (2004) Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain and Cognition 56:129–140.

Shenoy KV, Kao JC (2021) Measurement, manipulation and modeling of brain-wide neural population dynamics. Nat Commun 12:633.

Simon NW, Wood J, Moghaddam B (2015) Action-outcome relationships are represented differently by medial prefrontal and orbitofrontal cortex neurons during action execution. J Neurophysiol 114:3374–3385.

Smith PL, Ratcliff R (2004) Psychology and neurobiology of simple decisions. Trends in neurosciences 27:161–168.

Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD (2019) Distributed coding of choice, action and engagement across the mouse brain. Nature 576:266–273.

Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD (2019) Spontaneous behaviors drive multidimensional, brainwide activity. Science 364:255.

Sylte OC, Muysers H, Chen H-L, Bartos M, Sauer J-F (2024) Neuronal tuning to threat exposure remains stable in the mouse prefrontal cortex over multiple days. PLOS Biology 22:e3002475.

Thorndike EL (1898) Animal intelligence: An experimental study of the associative processes in animals. The Psychological Review: Monograph Supplements 2:i–109.

van Maanen L, Brown SD, Eichele T, Wagenmakers EJ, Ho T, Serences J, Forstmann BU (2011) Neural correlates of trial-to-trial fluctuations in response caution. J Neurosci 31:17488–17495.

Vong L, Ye C, Yang Z, Choi B, Chua S Jr, Lowell BB (2011) Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. Neuron 71:142–154.

Wang Q et al. (2020) The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell 181:936-953 e20.

Yee DM, Leng X, Shenhav A, Braver TS (2022) Aversive motivation and cognitive control. Neuroscience & Biobehavioral Reviews 133:104493.

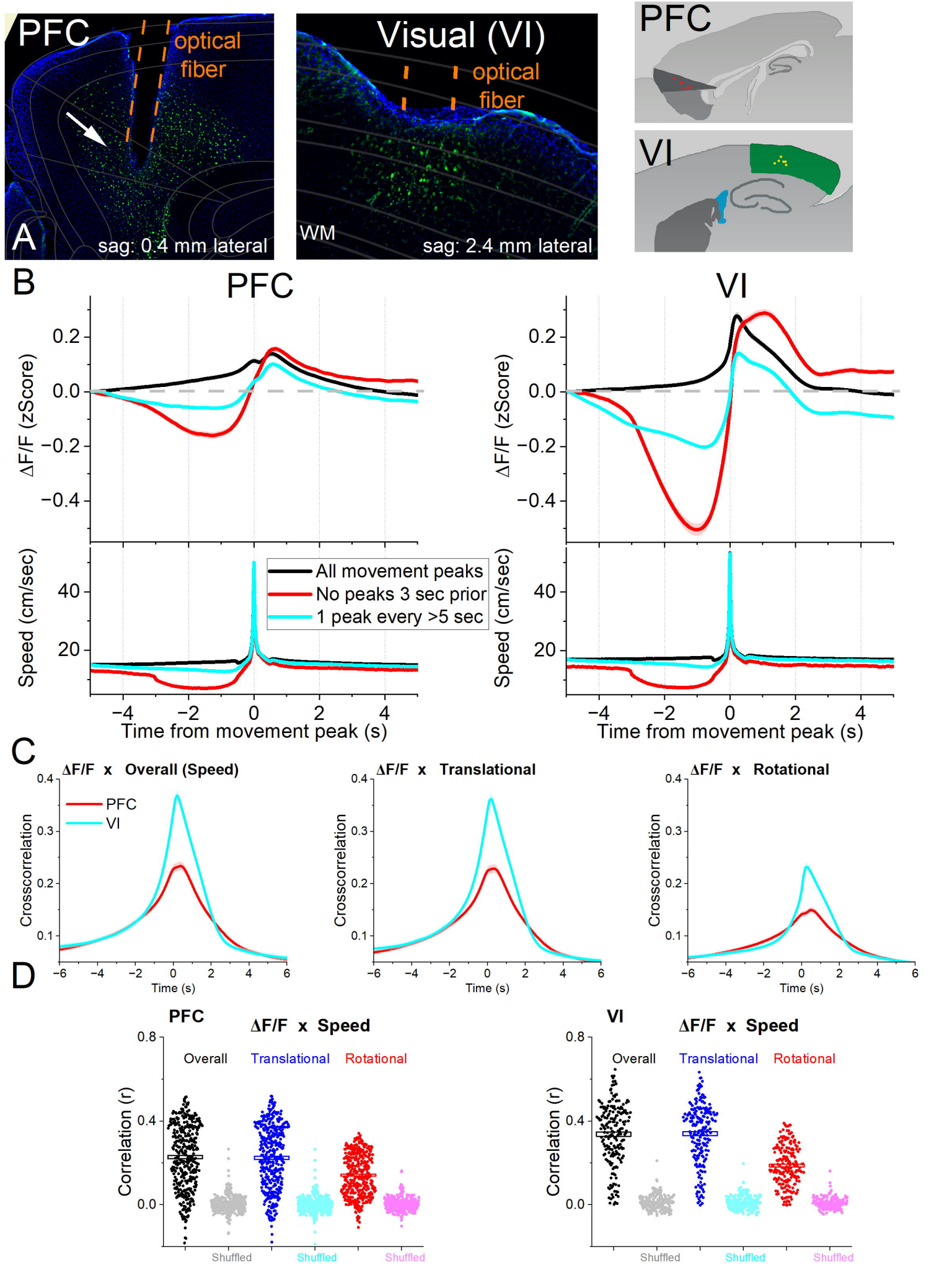
Zagha E, Erlich JC, Lee S, Lur G, O'Connor DH, Steinmetz NA, Stringer C, Yang H (2022) The Importance of Accounting for Movement When Relating Neuronal Activity to Sensory and Cognitive Processes. J Neurosci 42:1375–1382.

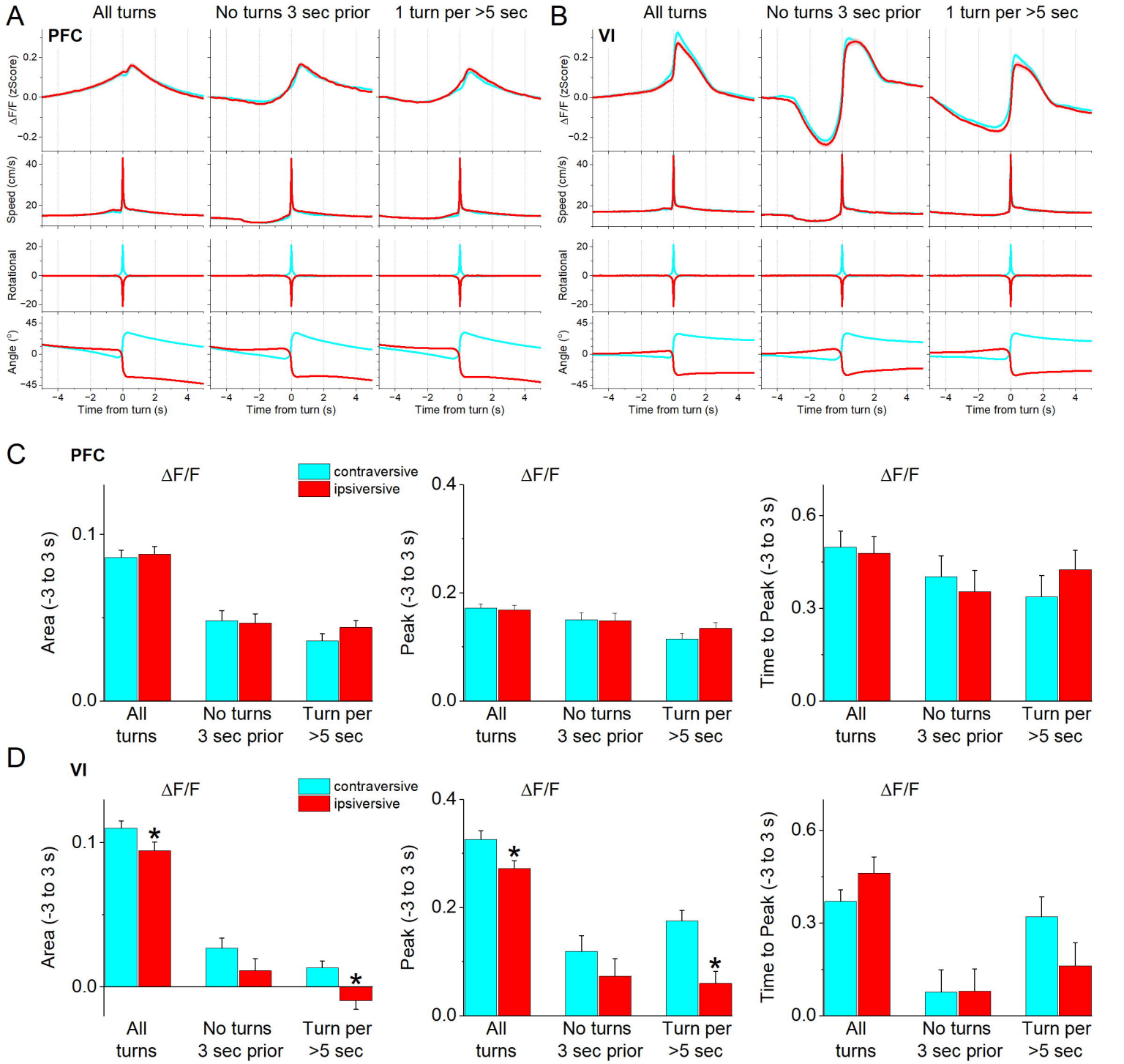
Zhou J, Hormigo S, Busel N, Castro-Alamancos MA (2023) The Orienting Reflex Reveals Behavioral States Set by Demanding Contexts: Role of the Superior Colliculus. J Neurosci 43:1778–1796.

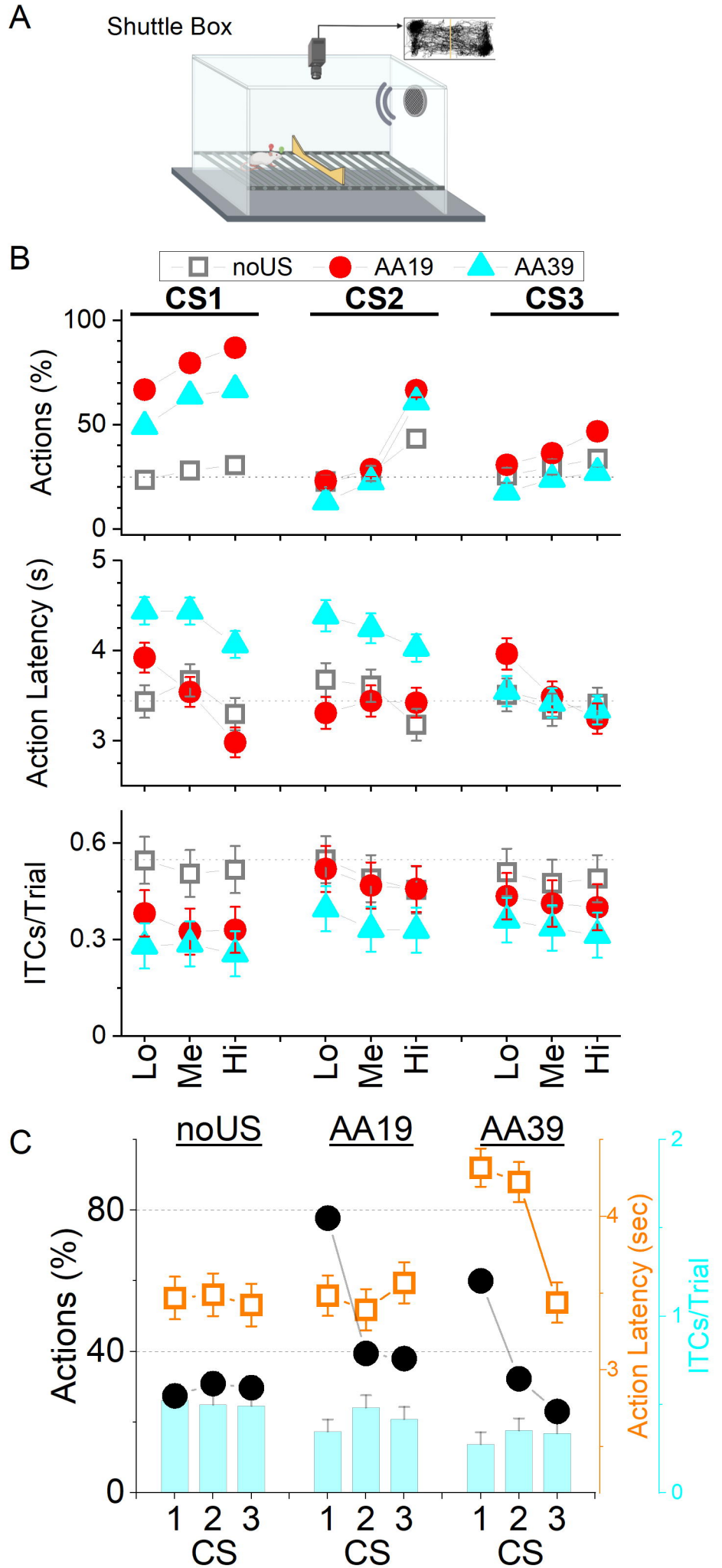
Zhou J, Hormigo S, Sajid MS, Castro-Alamancos MA (2022) Caution Influences Avoidance and Approach Behaviors Differently. J Neurosci 42:5899–5915.

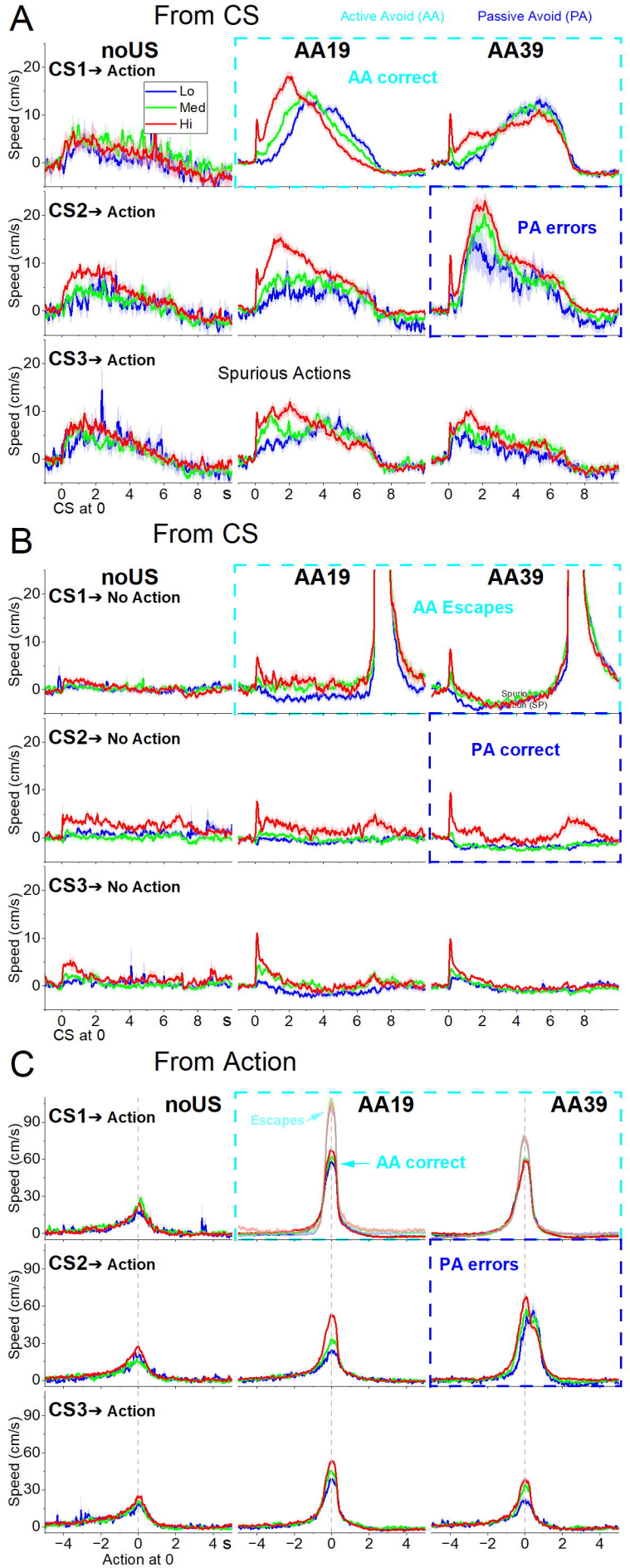
Zhou J, Hormigo S, Sajid MS, Castro-Alamancos MA (2024) Role of the Nucleus Accumbens in Signaled Avoidance Actions. eNeuro 11.

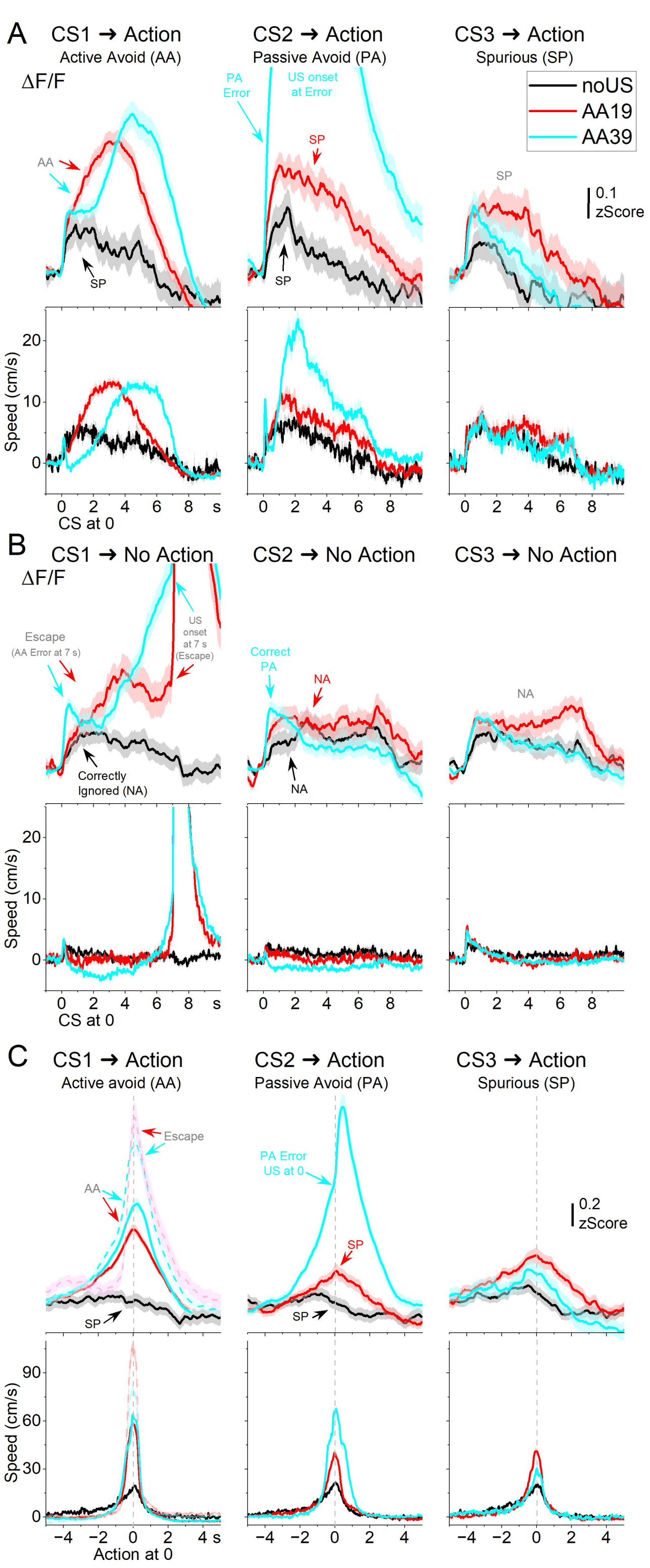
Zhou J, Sajid MS, Hormigo S, Castro-Alamancos MA (2025) A Forebrain Hub for Cautious Actions via the Midbrain. eLife 14.

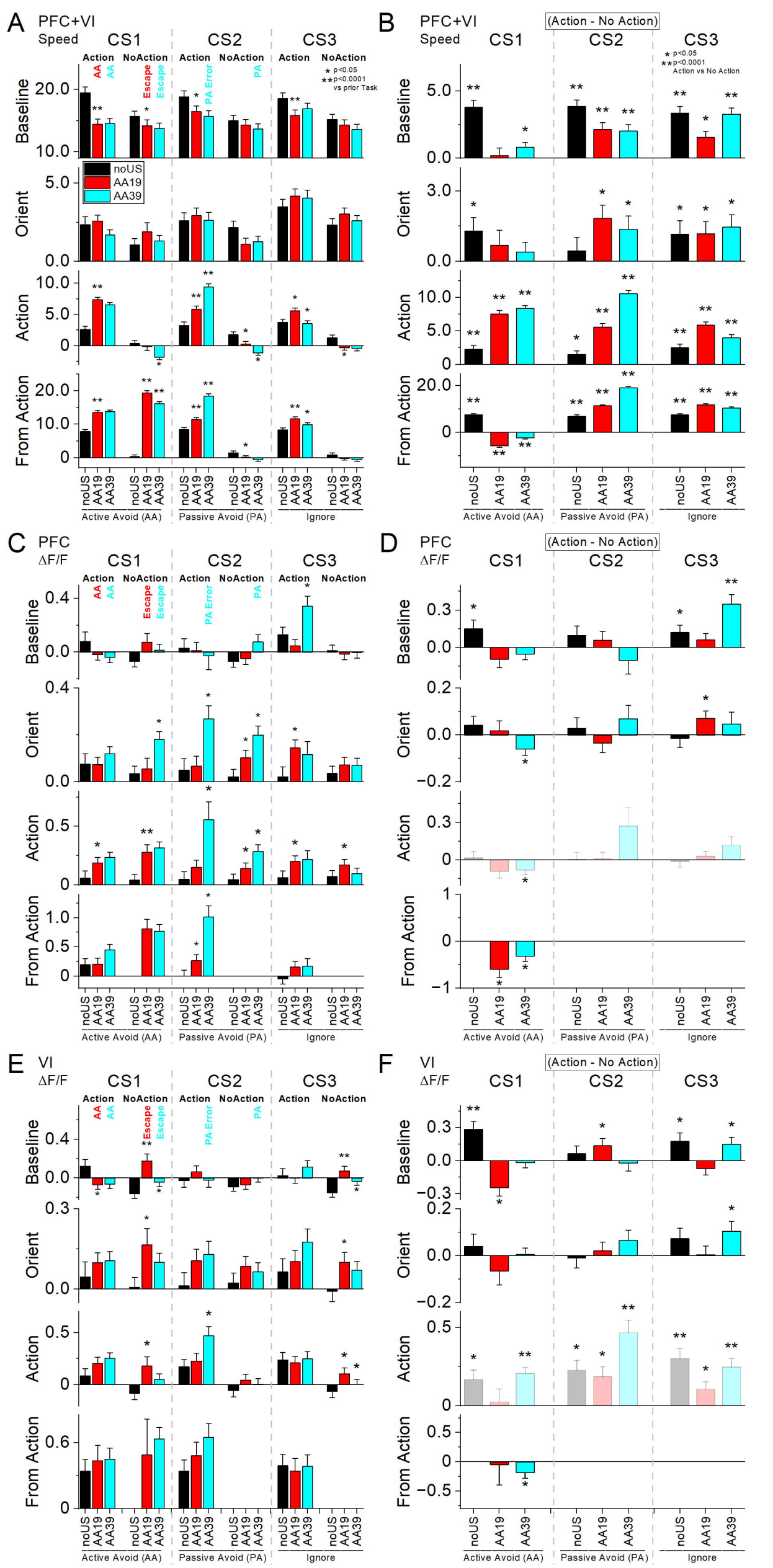

PFC
optical fiber
A
sag: 0.4 mm lateral
Visual (VI)
optical fiber
WM
sag: 2.4 mm lateral
PFC
VI
B
PFC
VI
ΔF/F (zScore)
Speed (cm/sec)
All movement peaks
No peaks 3 sec prior
1 peak every >5 sec
Time from movement peak (s)
C
ΔF/F x Overall (Speed)
ΔF/F x Translational
ΔF/F x Rotational
PFC
VI
Crosscorrelation
Time (s)
D
PFC
VI
ΔF/F x Speed
Overall
Translational
Rotational
Correlation (r)
Shuffled
Shuffled
Shuffled

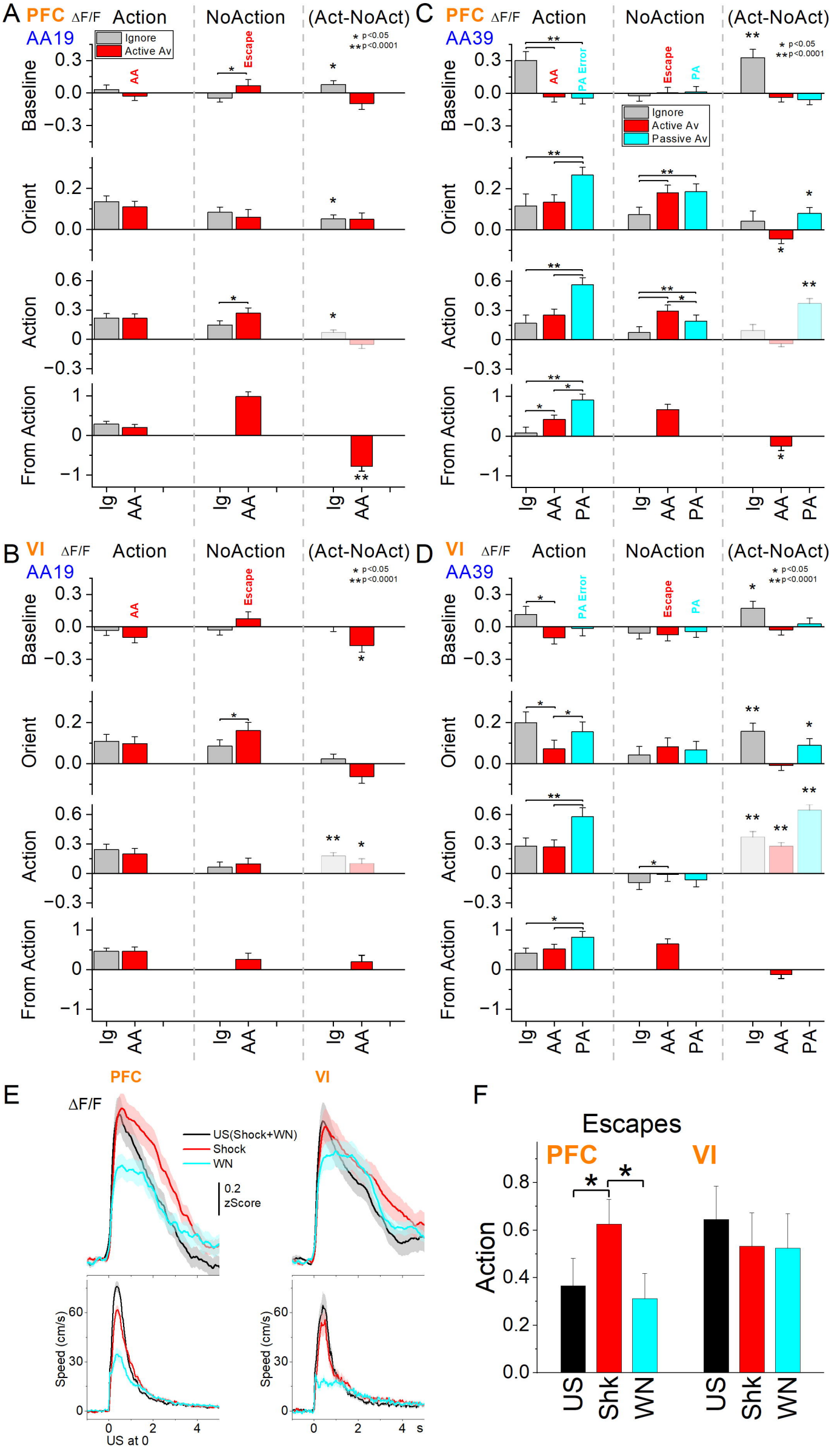

A
All turns
No turns 3 sec prior
1 turn per >5 sec
PFC
B
VI
ΔF/F (zScore)
Speed (cm/s)
Rotational
Angle (°)
Time from turn (s)
C
D
contraversive
ipsiversive
ΔF/F
Area (-3 to 3 s)
Peak (-3 to 3 s)
Time to Peak (-3 to 3 s)
All turns
No turns 3 sec prior
Turn per >5 sec
*

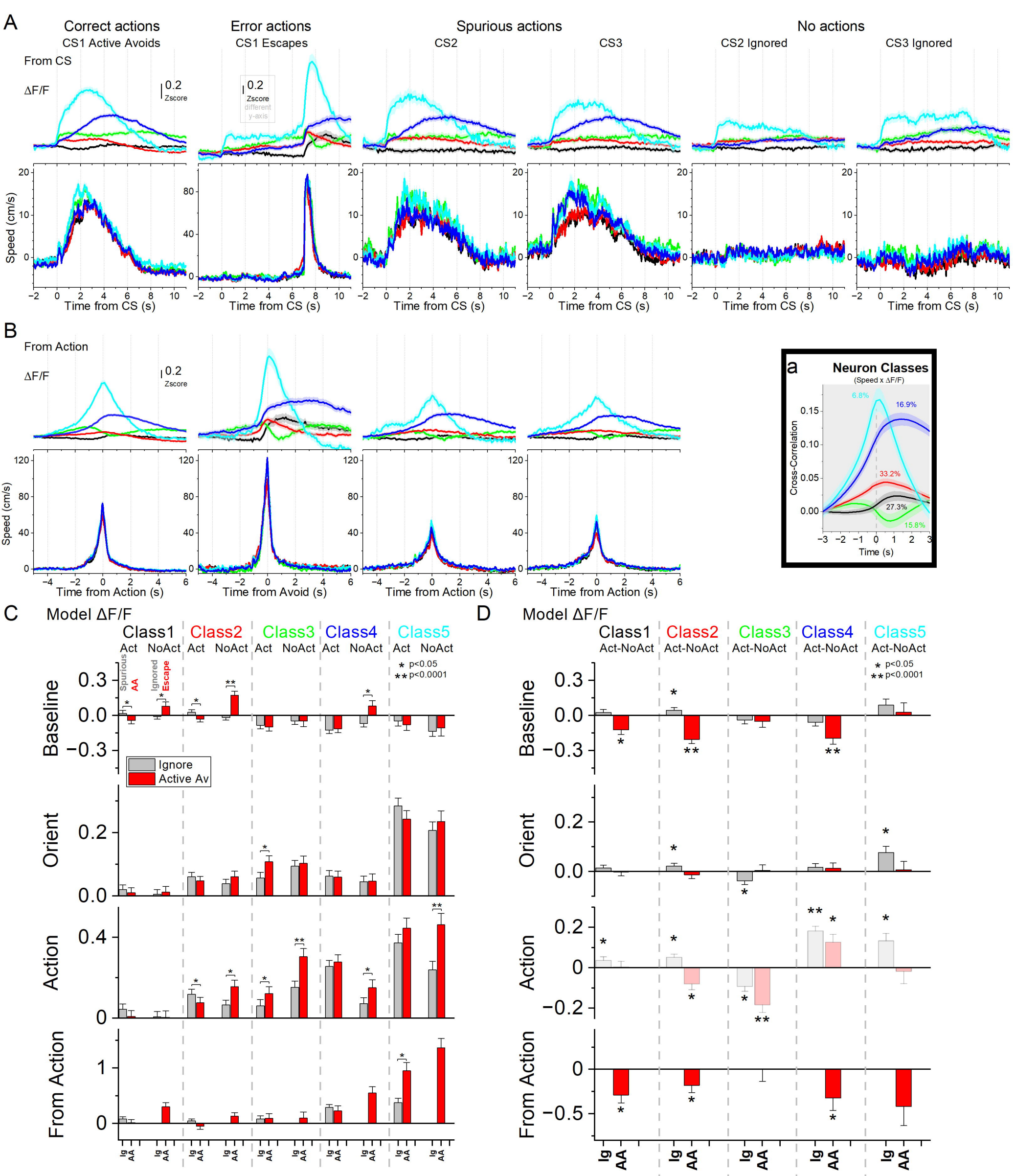

A
Shuttle Box
B
noUS
AA19
AA39
CS1
CS2
CS3
Actions (%)
Action Latency (s)
ITCs/Trial
Lo
Me
Hi
C
noUS
AA19
AA39
Actions (%)
Action Latency (sec)
ITCs/Trial
CS

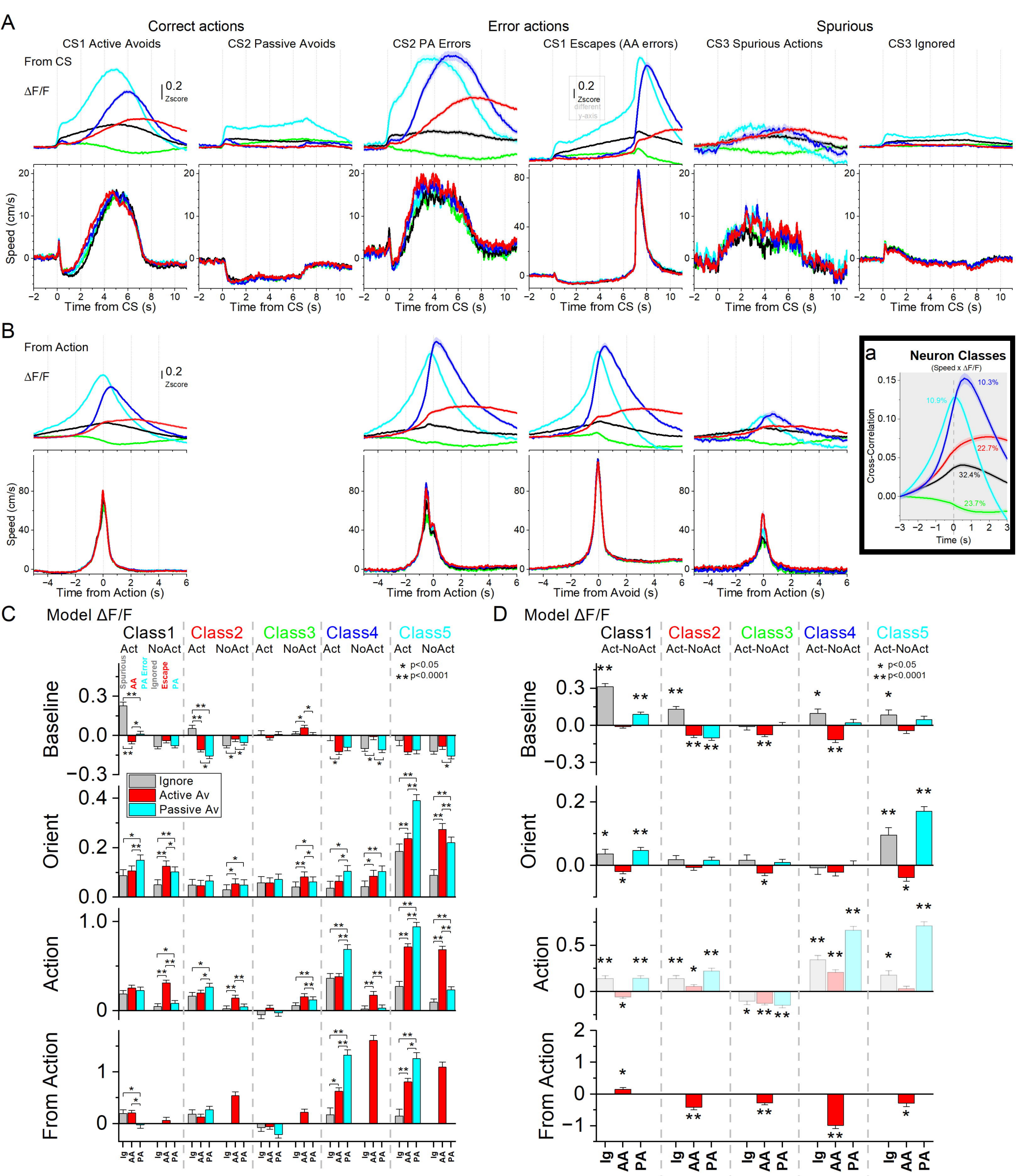

A
From CS
Active Avoid (AA)
Passive Avoid (PA)
noUS
AA19
AA39
CS1→ Action
Lo
Med
Hi
AA correct
CS2→ Action
PA errors
CS3→ Action
Spurious Actions
Speed (cm/s)
CS at 0
s
B
From CS
noUS
AA19
AA39
CS1→ No Action
AA Escapes
CS2→ No Action
PA correct
CS3→ No Action
Speed (cm/s)
CS at 0
s
C
From Action
noUS
AA19
AA39
CS1→ Action
Escapes
AA correct
CS2→ Action
PA errors
CS3→ Action
Speed (cm/s)
Action at 0
s

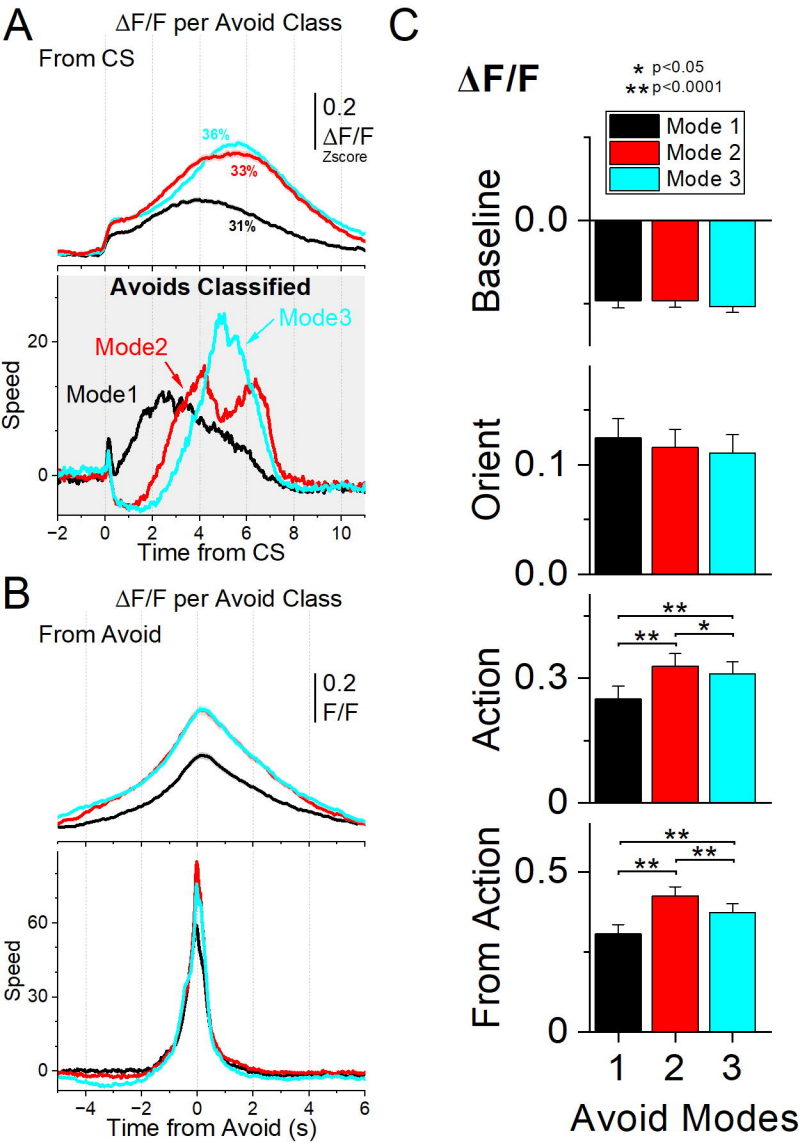

A
CS1 → Action
Active Avoid (AA)
CS2 → Action
Passive Avoid (PA)
CS3 → Action
Spurious (SP)
ΔF/F
PA Error
US onset at Error
noUS
AA19
AA39
AA
SP
0.1 zScore
Speed (cm/s)
CS at 0
B
CS1 → No Action
CS2 → No Action
CS3 → No Action
Escape (AA Error at 7 s)
US onset at 7 s (Escape)
Correct PA
NA
Correctly Ignored (NA)
C
Escape
PA Error US at 0
0.2 zScore
Action at 0

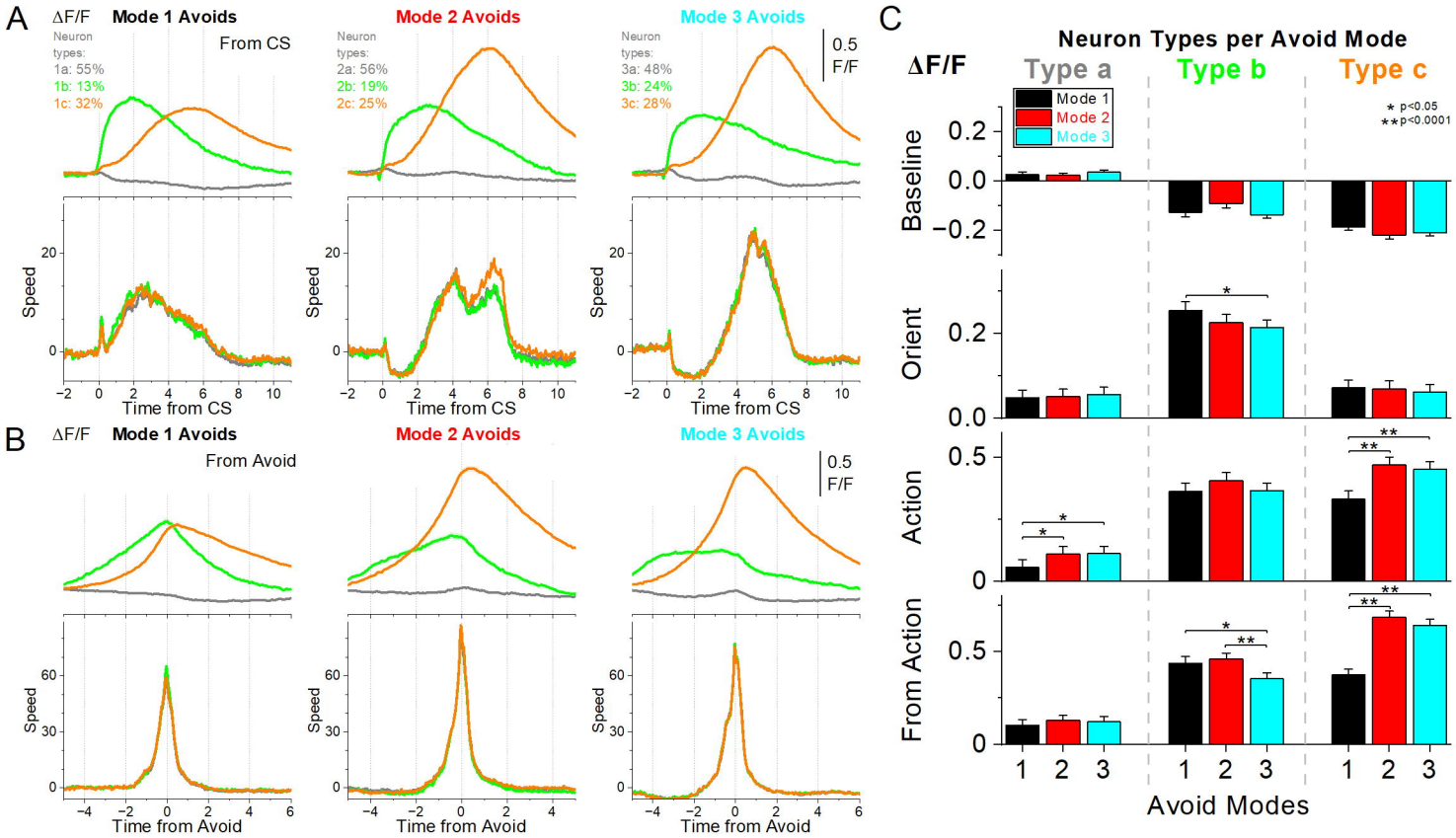

A
PFC+VI
Speed
CS1
CS2
CS3
Action
NoAction
AA
Escape
PA Error
PA
* p<0.05
** p<0.0001
vs prior Task
noUS
AA19
AA39
Baseline
Orient
Action
From Action
Active Avoid (AA)
Passive Avoid (PA)
Ignore
B
PFC+VI
(Action - No Action)
Action vs No Action
C
PFC
ΔF/F
D
E
VI
F

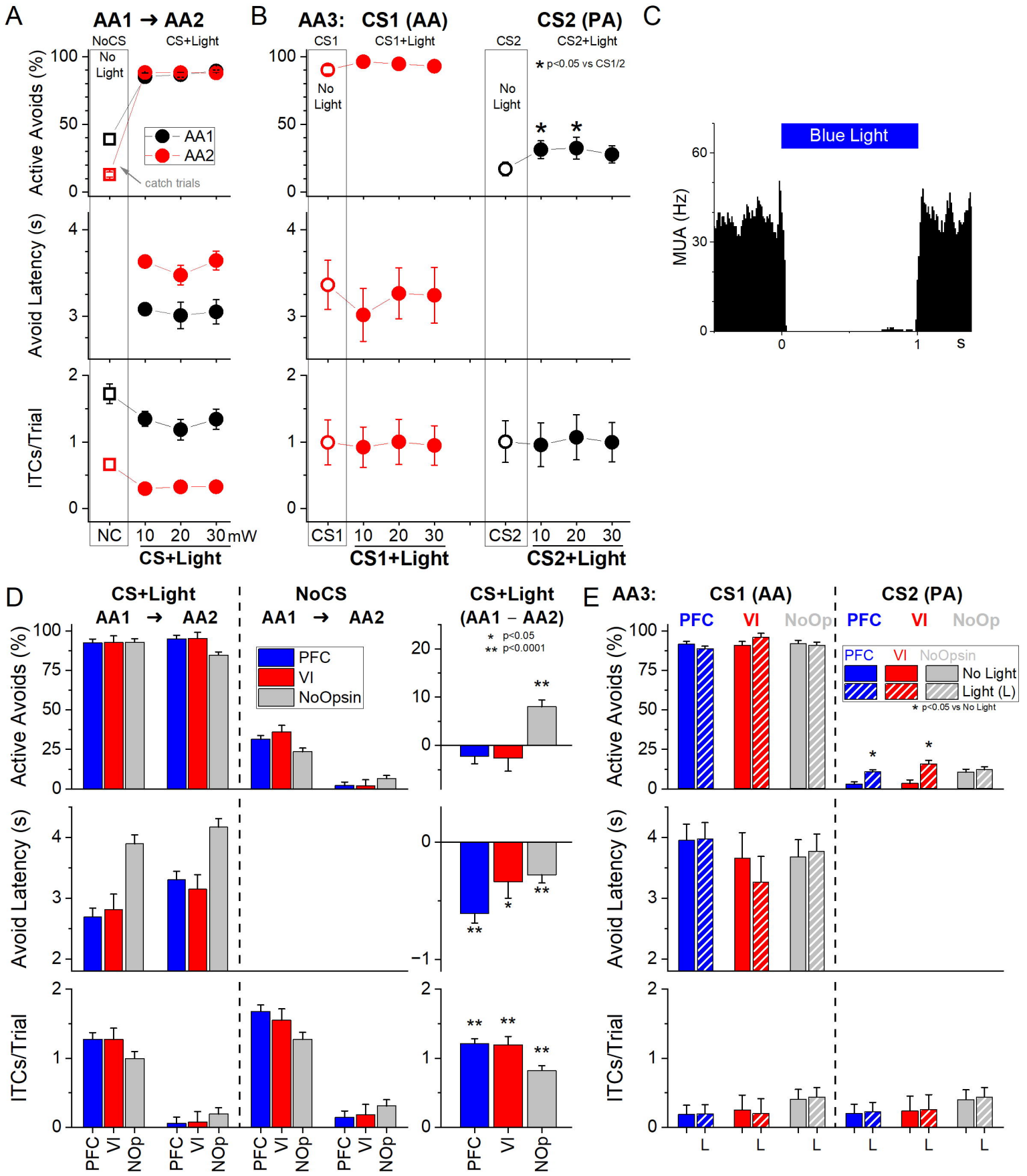

A
PFC ΔF/F
AA19
Action
NoAction
(Act-NoAct)
Ignore
Active Av
* p<0.05
**p<0.0001
Baseline
Orient
Action
From Action
AA
Escape
Ig
C
PFC ΔF/F
AA39
PA Error
PA
Passive Av
B
VI ΔF/F
D
VI ΔF/F
E
PFC
VI
ΔF/F
US(Shock+WN)
Shock
WN
0.2
zScore
Speed (cm/s)
US at 0
s
F
Escapes
PFC
VI
Action
US
Shk
WN


A
Correct actions
Error actions
Spurious actions
No actions
CS1 Active Avoids
CS1 Escapes
CS2
CS3
CS2 Ignored
CS3 Ignored
From CS
ΔF/F
0.2
Zscore
different y-axis
Speed (cm/s)
Time from CS (s)
B
From Action
Time from Action (s)
Time from Avoid (s)
a
Neuron Classes
(Speed x ΔF/F)
Cross-Correlation
Time (s)
6.8%
16.9%
33.2%
27.3%
15.8%
C
Model ΔF/F
Class1
Class2
Class3
Class4
Class5
Act
NoAct
Spurious
AA
Ignored
Escape
* p<0.05
** p<0.0001
Baseline
Orient
Action
From Action
Ignore
Active Av
Ig
AA
D
Act-NoAct


A
Correct actions
Error actions
Spurious
CS1 Active Avoids
CS2 Passive Avoids
CS2 PA Errors
CS1 Escapes (AA errors)
CS3 Spurious Actions
CS3 Ignored
From CS
ΔF/F
0.2 Zscore
Speed (cm/s)
Time from CS (s)
B
From Action
Time from Action (s)
Time from Avoid (s)
a
Neuron Classes
(Speed x ΔF/F)
Cross-Correlation
Time (s)
10.3%
10.9%
22.7%
32.4%
23.7%
C
Model ΔF/F
Class1
Class2
Class3
Class4
Class5
Act
NoAct
Spurious
AA
PA Error
Ignored
Escape
PA
* p<0.05
** p<0.0001
Baseline
Orient
Action
From Action
Ignore
Active Av
Passive Av
Ig
AA
PA
D
Model ΔF/F
Act-NoAct


A
ΔF/F per Avoid Class
From CS
0.2
ΔF/F
Zscore
36%
33%
31%
Avoids Classified
Mode3
Mode2
Mode1
Speed
20
0
−2
0
2
4
6
8
10
Time from CS
B
ΔF/F per Avoid Class
From Avoid
0.2
F/F
Speed
60
30
0
−4
−2
0
2
4
6
Time from Avoid (s)
C
ΔF/F
* p<0.05
** p<0.0001
Mode 1
Mode 2
Mode 3
Baseline
0.0
Orient
0.1
0.0
Action
0.3
0
From Action
0.5
0
**
*
1
2
3
Avoid Modes


A
ΔF/F
Mode 1 Avoids
From CS
Neuron types:
1a: 55%
1b: 13%
1c: 32%
Mode 2 Avoids
Neuron types:
2a: 56%
2b: 19%
2c: 25%
Mode 3 Avoids
Neuron types:
3a: 48%
3b: 24%
3c: 28%
0.5 F/F
Speed
Time from CS
B
ΔF/F
Mode 1 Avoids
From Avoid
Mode 2 Avoids
Mode 3 Avoids
0.5 F/F
Speed
Time from Avoid
C
Neuron Types per Avoid Mode
ΔF/F
Type a
Type b
Type c
Mode 1
Mode 2
Mode 3
* p<0.05
** p<0.0001
Baseline
Orient
Action
From Action
Avoid Modes


A
AA1 → AA2
NoCS
CS+Light
No Light
Active Avoids (%)
AA1
AA2
catch trials
Avoid Latency (s)
ITCs/Trial
NC
10 20 30 mW
CS+Light
B
AA3: CS1 (AA)
CS2 (PA)
CS1
CS1+Light
CS2
CS2+Light
* p<0.05 vs CS1/2
No Light
CS1
10 20 30
CS1+Light
CS2
10 20 30
CS2+Light
C
Blue Light
MUA (Hz)
s
D
CS+Light
AA1 → AA2
NoCS
AA1 → AA2
CS+Light
(AA1 – AA2)
* p<0.05
** p<0.0001
PFC
VI
NoOpsin
Active Avoids (%)
Avoid Latency (s)
ITCs/Trial
PFC VI NOp PFC VI NOp PFC VI NOp PFC VI NOp PFC VI NOp
E
AA3:
CS1 (AA)
CS2 (PA)
PFC VI NoOp PFC VI NoOp
PFC VI NoOpsin
No Light
Light (L)
* p<0.05 vs No Light
Active Avoids (%)
Avoid Latency (s)
ITCs/Trial
L L L L L L