Neural Representation of Associative Threat Learning in Pulvinar Divisions, Lateral Geniculate Nucleus, and Mediodorsal Thalamus in Humans

Muhammad Badarnee[1], Zhenfu Wen[1], B. Isabel Moallem[1], Stephen Maren[2,3], and Mohammed R Milad[1*]

[1]Department of Psychiatry and Behavioral Sciences, The University of Texas, Health Science Center at Houston, McGovern Medical School, TX USA.

[2]Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, IL

[3]Department of Psychology, University of Illinois Urbana-Champaign, Champaign, IL

***Corresponding Author:**

Mohammed R Milad, PhD

Department of Psychiatry and Behavioral Sciences

UTHealth, McGovern Medical School

1941 East Road, BBSB - 2118

Houston, Texas 77054

mohammed.r.milad@uth.tmc.edu

Abstract

Understanding the neural mechanisms underlying associative threat learning is essential for advancing behavioral models of threat and adaptation. We investigated distinct activation patterns across thalamic pulvinar divisions, lateral geniculate nucleus (LGN), and mediodorsal thalamus (MD) during the acquisition of associative threat learning in the MRI. The anterior pulvinar and MD exhibited parallel activation patterns, which we interpret as relating to automatic and more deliberative learning processes. Additionally, our findings suggest a hierarchical pulvinar organization during fear conditioning, in which coordinated activation among inferior, lateral, medial, and anterior divisions may support the integration of threat-related information. Pulvinar divisions and the MD showed activation during extinction learning and exhibited patterns consistent with salience processing and safety–threat memory expression during extinction recall and threat renewal. LGN activation patterns during threat learning were consistent with feedforward processing of visual information. This study extends dominant brain models of threat learning and memory, reframing our understanding of distinct thalamic roles in these psychological processes.

Keywords: Thalamus; Pulvinar; Lateral geniculate nucleus; Mediodorsal thalamus; Threat learning; Conditioning; Extinction; Recall extinction; Threat renewal; Emotional memory.

Introduction

Over a century ago, Ivan Pavlov provided foundational behavioral evidence for associative learning, demonstrating that pairing a neutral stimulus with an unconditioned stimulus could elicit a conditioned response (Pavlov, 1904). This discovery laid the groundwork for modern learning theories and has been particularly influential in understanding how humans learn to associate neutral stimuli with threats, a process known as associative threat learning. This survival-oriented learning plays a key role in shaping a broad range of behavioral and emotional responses, such as avoidance, decision-making under threats, and fear regulation (Badarnee et al., 2025; Kolling et al., 2014; Korn and Bach, 2019, 2018; Milad et al., 2013). Research on the neural mechanisms underlying associative threat learning has primarily focused on brain regions involved in emotional regulation, such as the amygdala, hippocampus, and prefrontal cortex (PFC) (Fullana et al., 2015; Krasne et al., 2021; Milad and Quirk, 2012; Vuilleumier et al., 2003). More recently, the thalamus has received growing attention in threat learning (Lithari et al., 2015; Penzo et al., 2015; Ramanathan et al., 2018; Ramanathan and Maren, 2019; Ratigan et al., 2023; Totty et al., 2023), since its role has been reconsidered beyond the traditional view of a mere relay station. Yet, the distinct contribution of this complex structure to associative threat learning remains poorly understood.

The thalamus is a hub structure in the mammalian brain that plays multiple critical roles, ranging from basic sensory processing to higher-order functions and threat learning (Halassa and Sherman, 2019; Hummos et al., 2022; Hwang et al., 2017; Saalmann and

Kastner, 2011; Sherman, 2016, 2007). It has long been proposed that threat detection is mediated by the pulvinar and the lateral geniculate nucleus (LGN) (Carr, 2015; Pessoa and Adolphs, 2010; Silverstein and Ingvar, 2015), two major nuclei of the visual thalamus (Arcaro et al., 2015; Casanova and Chalupa, 2023; Takakuwa et al., 2021). The pulvinar projects directly to the amygdala and is believed to subconsciously facilitate rapid survival reactions to threats via a subcortical path ('low road') (Carr, 2015; Pessoa and Adolphs, 2010; Rafal et al., 2015; Wei et al., 2015). The LGN, on the other hand, is part of a slower but more accurate neural pathway, connecting peripheral information to cortical regions for comprehensive processing ('high road'). Additionally, the involvement of the mediodorsal thalamus (MD) in threat learning has also been proposed(Lee and Shin, 2016; Lee et al., 2011). Although this nucleus is not typically classified as part of the visual thalamus, it serves as a high-order region reciprocally connected to the amygdala and PFC while playing a critical role in executive functions (Hwang et al., 2020; Li et al., 2022; Mukherjee et al., 2021; Wolff and Halassa, 2024).

The general implication of these nuclei in threat processing has been demonstrated in primates and rodent models. It has been shown that exposing primates to snake images elicited increased neuronal firing within the pulvinar (Le et al., 2013). Excitation of neurons within the LGN enhanced the acquisition of eyeblink conditioning in rodents (Halverson and Freeman, 2010). Inhibiting this nucleus was associated with impaired conditioned responses (Shi and Davis, 2001; Steinmetz et al., 2013), and activating GABA neurons in the LGN reduced freezing response to an overhead dark shadow that mimics a real-world predator in rats (Salay and Huberman, 2021). Reduced freezing

was also observed after MD lesions (Li et al., 2004) and when the connections with the anterior cingulate cortex were ablated (Zheng et al., 2020). Lesions within the MD are also associated with impaired fear extinction (Lee et al., 2011), and injecting gabazine, a modulator of extra-synaptic GABA receptors, into the MD facilitated extinction (Paydar et al., 2014).

Evidence regarding the involvement of the thalamus in the context of threat learning in humans is relatively sparse. Much of our knowledge comes from studies focusing on attention, perception, and decision-making. For example, the LGN has been mainly viewed as a transmitter of peripheral information to other brain regions and found to be associated with selective attention and anticipation of visual stimuli (Mahoney and Schmidt, 2024; O'Connor et al., 2002; Saalmann and Kastner, 2009). The MD contribution to perception, spatial attention, and decision-making has also been reported (Griffiths et al., 2022; Wurtz et al., 2011). Beyond cognitive research, some human studies have specifically investigated the pulvinar's engagement in threat detection. Individuals with increased fiber density in the pulvinar-amygdala pathway showed an enhanced ability to recognize fearful faces (McFadyen et al., 2019). A patient with a complete lesion in the left pulvinar showed slower responses to threatening images when stimuli were presented on the ipsilesional, but not contralesional, field (Ward et al., 2005). In agreement with this, unseen stimuli (fearful vs. happy faces) presented on the blind side of patients with hemianopia moderated their performance when the pulvinar was spared but not when it was lesioned (Bertini et al., 2018).

Specific thalamic contribution to threat processing is still largely unknown, and translational research on this topic is limited. We aim to address these critical gaps in our understanding by investigating the distinct neural representations of associative threat learning in the human pulvinar divisions, LGN, and MD. We analyzed the neuroimaging data of 293-412 controls. All participants underwent a two-day threat learning paradigm (Milad et al., 2009, 2007; Wen et al., 2024, 2022) while in the fMRI scanner. The conditioned stimulus (CS+) (e.g., red light) was associated with an electric shock (unconditioned stimulus US, 62.5% reinforcement rate), while a control light (e.g., blue) was never paired with the US (CS-). This conditioning phase occurred in a computer-displayed visual context A (e.g., an office). The extinction learning included presenting the CS+ and CS- with no US reinforcement in a distinct context B (e.g., bookcase). In extinction recall and threat renewal, the extinguished stimuli were presented with no US reinforcement within contexts B (safe contextual cues) and A (threat contextual cues), respectively. Schematic illustrations of each phase of the paradigm are presented in **Figs. 1a, 4a-6a**.

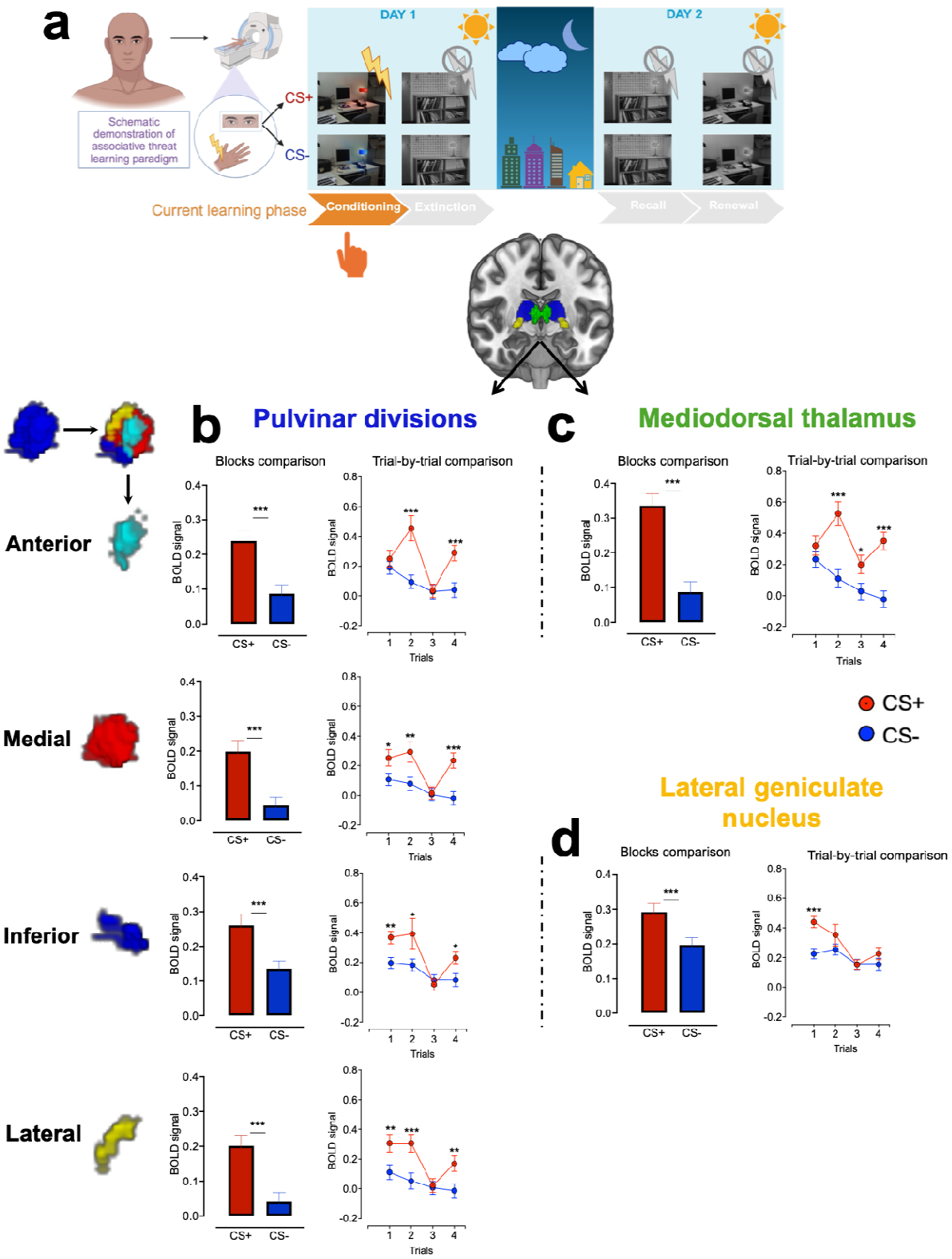

Fig. 1. Neural representation during associative threat learning in pulvinar divisions, MD, and LGN.
(a) Human fear conditioning paradigm during fMRI. (b–d) Means ± SE of activation in response to

CS+ vs. CS− at both block-wise and trial-wise levels within pulvinar divisions (b), MD (c), and LGN (d). Block-level comparisons were assessed using paired *t*-tests, while trial-level effects were examined using a 2 × 2 repeated-measures ANOVA, followed by post hoc comparisons between CS+ and CS− across the four trials. Multiple comparisons were controlled using false discovery rate (FDR) correction. Conditioning sample size: $n = 293$.
Detailed statistical parameters are provided in **Supplementary Tables 1–2**. Additionally, cross-validation analyses were conducted to assess the robustness of the CS+ vs. CS− signal in the anterior pulvinar and MD, consistent with our interpretation of learning-related effects (**Supplementary Figs. 1–2** and **Supplementary File 1**).

*p < 0.05, **p < 0.01, ***p < 0.001. SE, standard error; MD, mediodorsal thalamus; LGN, lateral geniculate nucleus; CS+, conditioned stimulus predicting shock; CS−, conditioned stimulus predicting no shock. Display items in panel (a) were created using BioRender (BioRender.com/t54zvyf).

We focus on neural activation patterns within pulvinar divisions, LGN, and MD while acquiring the CS-US association. As associative learning is a rapid psychological process (Konrad et al., 2024), we analyzed the first four trials of all threat learning phases. The purpose of this study is to investigate thalamic function during each learning phase separately, focusing on CS+ vs. CS− differences within phases rather than comparing activation across phases. This phase-specific approach allows us to characterize thalamic functional dynamics within each stage of learning and memory, avoiding potential confounds arising from the distinct processes of conditioning, extinction, and recall. We compared brain activation (BOLD) to the CS+ vs. CS- at the block level by averaging the activation across all four trials and at a trial-by-trial level to provide finer activation temporal resolution. This dual approach allows us to capture the general neural responses associated with associative threat learning and provides new insights into trial-level dynamics (Wen et al., 2022). The current dominant neurocognitive thalamic models (Sherman, 2007; Sherman and Guillery, 2006) highlight the thalamus's role in mediating cortical-cortical communications and facilitating high-order functions. Based on this view, we anticipate distinct pulvinar, MD, and LGN roles.

The LGN, as a first-order nucleus, is expected to serve relay functions, transmitting information with minimal integrations. The pulvinar and MD, on the other hand, as higher-order nuclei, are likely involved in more complex processing, potentially integrating threat-related information.

Results

Parallel functional representation consistent with associative threat learning in the anterior pulvinar and MD during conditioning

At the block level, both the anterior pulvinar and MD showed increased activation to CS+ vs. CS− (anterior pulvinar: $t_{(292)} = 4.41$, $p = 0.00001$, $d = 0.25$; MD: $t_{(292)} = 6.41$, $p = 5.83x10^{-10}$, $d = 0.37$; **Fig. 1b–c**), suggesting a possible involvement of these regions in early associative threat learning. The trial-wise analysis revealed a distinct activation pattern. On the first trial, activation did not differ between CS+ and CS− in either the anterior pulvinar ($p_{FDR} = 0.44$) or the MD ($p_{FDR} = 0.20$). In contrast, on the subsequent trial, both regions showed significantly greater BOLD responses to CS+ than CS− (anterior pulvinar: $p_{FDR} = 0.000001$; MD: $p_{FDR} = 0.000003$; **Fig. 1b–c**; Detailed statistical parameters are provided in **Supplementary Tables 1–2**). In our paradigm, the electric shock was paired with CS+ at the end of the CS presentation. The similarity in BOLD responses to CS+ and CS- during the first trial likely reflects an initial equivalence in the emotional valence of the stimuli. The gradual increase in activation for CS+ by the second trial suggests a shift in the emotional valence, consistent with rapid associative learning.

The similarity in trial-level activation patterns in the anterior pulvinar and MD raises the question of whether this apparent similarity truly associated with a parallel functional contribution to associative threat learning. To test this possibility, we first quantified the trial-wise relationships using Pearson correlation coefficients between the two regions. The observed correlations support consistent co-activation across all corresponding trials (*r values* ≥ *0.66, p* < *0.001*; **Fig. 2a**, left). However, this approach could not exclude the possibility that the shared variance is a consequence of shared anatomical proximity, particularly because both regions are part of the same brain structure, i.e., the thalamus. To control for this confound, we conducted a hierarchical regression model in four steps for each trial. We modeled MD activation as the target and progressively added the pulvinar divisions as covariates. Despite minor variations in regression coefficients across control models, the anterior pulvinar consistently showed the largest standardized regression coefficients (all $p < 0.01$; **Fig. 2a**, middle) and accounted for a substantial proportion of the explained variance in MD activation ($R^2 = 0.43–0.61$). In contrast, other pulvinar divisions contributed minimally to the explained variance (*across all other pulvinar divisions, the maximum observed* ΔR^2 *was 0.02* in trial 1, *0.01* in trial 2, *0.01* in trial 3, and *0.002* in trial 4; **Fig. 2a**). These results were stable across 10,000 bootstrap resamples with replacement, indicating that the association between anterior pulvinar activation and MD activation is robust to resampling variability (**Fig. 2a**, right; additionally, see alternative hypothesis tests of the anterior pulvinar–MD relationships in **Supplementary Fig. 3**).

Applying the same analytical approach while additionally controlling for LGN activation as an anatomical control, beyond the pulvinar itself, yielded comparable results. Anterior

pulvinar activation continued to account for the majority of the explained variance in MD activation, whereas the inclusion of LGN activation produced only an incremental additional increase in explained variance ($\Delta R^2 = 0.006$ in trial 1, *0.01* in trial 2, *0.03* in trial 3, and *0.01* in trial 4; **Fig. 2b**). Together, these analyses support preferential trial-wise functional co-activations between the anterior pulvinar and MD, independent of shared anatomical proximity.

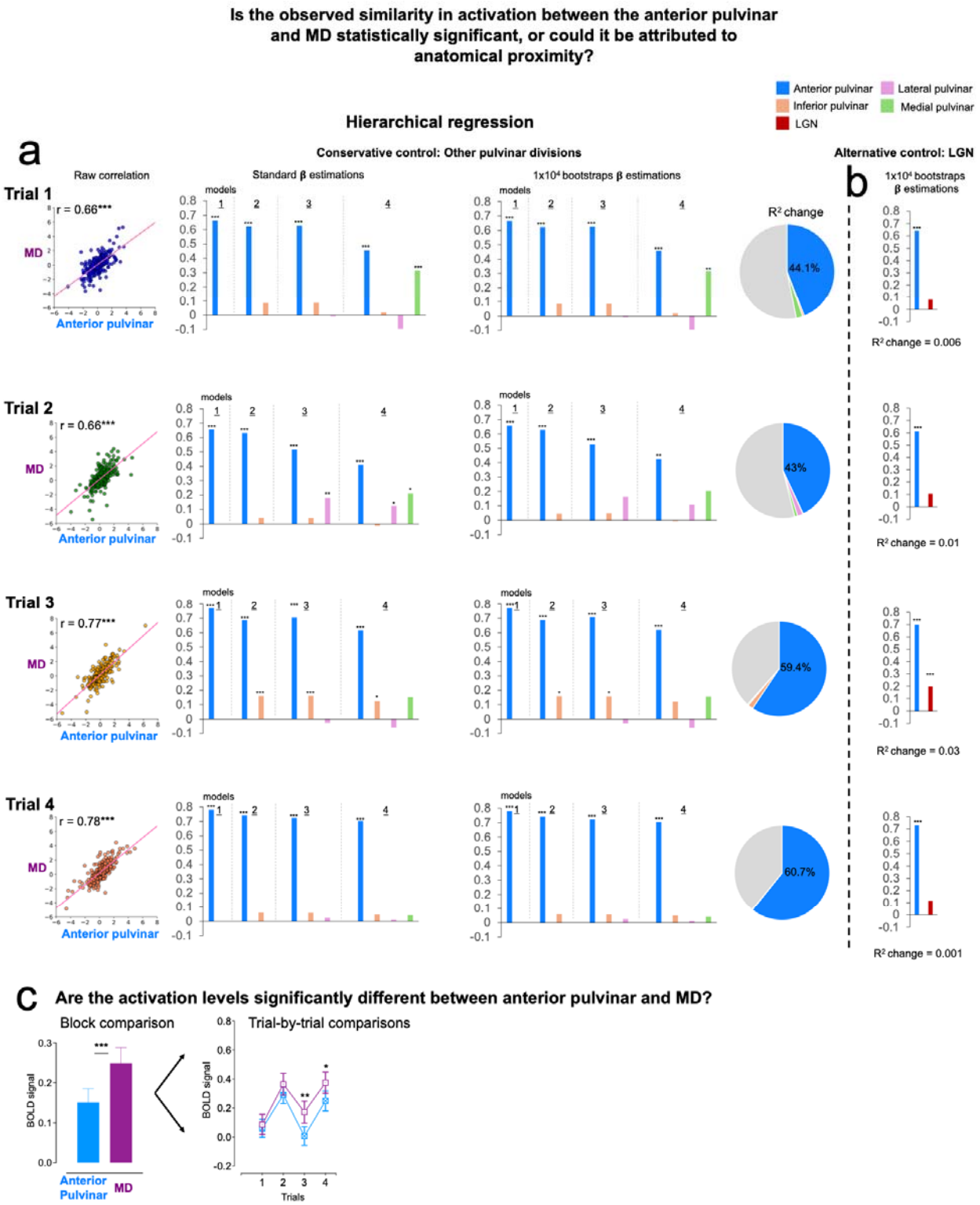

Fig 2. Quantifying the relationships between the anterior pulvinar and MD during conditioning. a. Hierarchical regression models of trial-wise relationships between the anterior pulvinar and MD activations while controlling for a potential effect of anatomical proximity. The effect of anatomical proximity was controlled by progressively adding other pulvinar divisions as controls. **b.** Results of a hierarchical model that included the LGN as an alternative control, distinct from the

pulvinar. **c.** Comparison of activation levels in the anterior pulvinar and MD at both block-wise and trial-wise levels (means ± SE). Sample size in these analyses: n = 293. Supplementary analyses testing alternative hypotheses regarding the correlations between the MD and all pulvinar divisions are presented in **Supplementary Fig. 3**. *p<0.05, **p<0.01, ***p<0.001 MD: Mediodorsal thalamus. LGN: Lateral geniculate nucleus. SE: Standard error.

Although co-activation does not necessarily imply similar activation magnitude, we further tested the differences in activation levels in the anterior pulvinar and MD. We used a t-test to capture the differences in the overall activation at the block level and repeated measures analysis of variance (RM-ANOVA) to capture the activation differences at the trial level. The results showed that the overall MD activation in response to CS+ was higher than the anterior pulvinar response ($t_{(292)} = 3.49$, $p = 0.0005$, $d = 0.20$, **Fig. 2c**). -The trial-wise analysis revealed that the overall differences between the two regions is driven particularly from the activation in trials 3 and 4 ($F_{(1, 292)} = 5.84$, $p = 0.01$; Trial 1: $t_{(292)} = 0.49$, $p_{FDR} = 0.82$, $d = 0.02$; Trial 2: $t_{(292)} = 0.03$, $p_{FDR} = 0.97$, $d = 0.002$, Trial 3: $t_{(292)} = 3.40$, $p_{FDR} = 0.003$, $d = 0.199$; Trial 4: $t_{(292)} = 2.63$, $p_{FDR} = 0.01$, $d = 0.15$; **Fig. 2c**). The increased activation within the MD might be associated with different levels of threat processing in the two regions.

A data-driven approach reveals hierarchical functional processing of threat in pulvinar divisions

The anatomical and functional interconnections between pulvinar divisions remain insufficiently characterized. To address this gap and understand the pulvinar's role in early associative threat learning, we integrated findings from complementary analyses into a novel computational model designed to investigate the functional relationships between pulvinar divisions.

We first defined the activation within each pulvinar division as the differences in BOLD signal at the block level in response to CS+ compared to CS-. We then performed RM-ANOVA to test whether the pulvinar divisions were engaged with different activation levels. We found no activation differences, suggesting similar processing levels of CS+ information across all pulvinar divisions ($F_{(2.29,\ 670.33)} = 0.96,\ p = 0.39$; **Fig. 3a**). Using network analysis, we explored the underlying functional dynamics between the divisions. We employed the EBICglasso method to estimate a sparse Gaussian graphical model (Foygel and Drton, 2010). To assess the stability of the network and provide robust estimates of edges and centrality measures, we conducted 10,000 bootstrap resamples. This approach allowed us to compute 95% confidence intervals (CI), ensuring the reliability of our findings. The combined use of EBICglasso and bootstrapping is essential for accurately capturing the dynamics of pulvinar division interactions. The resulting network included four nodes and five edges. The network plot pointed to stable direct edges connecting the lateral pulvinar to the anterior pulvinar and indirect edges connecting both the lateral and inferior pulvinar to the anterior through the medial pulvinar (**Fig. 3b**). Additionally, the medial pulvinar exhibited the highest centrality values, a measure that captures the node importance within a network, indicating a potential hub role of this division in associative threat learning (**Fig. 3b**).

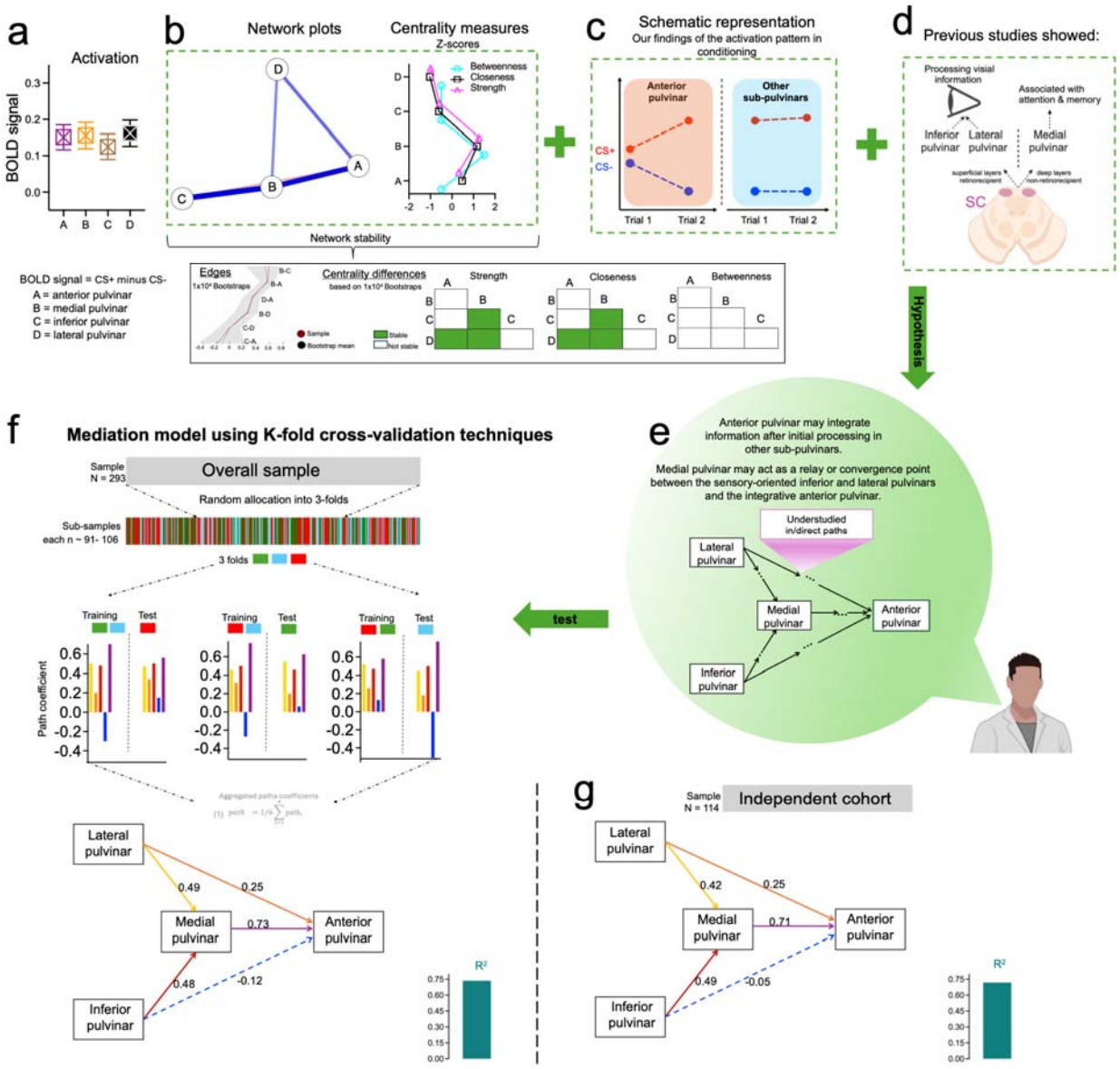

Fig 3. A data-driven approach to understanding the functional relationships between pulvinar divisions during conditioning. a. Means ± SE of activation differences between pulvinar divisions. **b.** Network analysis reveals that the medial pulvinar serves as a central hub, mediating interactions among pulvinar divisions and exhibiting increased centrality measures. **c.** Schematic visualization of activation patterns we observed in pulvinar divisions. **d.** Previous studies suggest that the inferior and lateral pulvinar are involved in processing basic visual information (Berman and Wurtz, 2010, 2008; Cortes et al., 2024), while the medial pulvinar is associated with higher-level functions, including working memory (Homman-Ludiye and Bourne, 2019). **e.** Based on **b-d**, we hypothesize that the medial pulvinar mediates the relationships with other divisions. **f.** Mediation analysis supports our hypothesis (panel **e**). **g.** Validation of the mediation model on an additional independent sample. Dashed paths in panels **f** and **g** represent statistically unstable paths, while the continuous paths indicate stable paths. *p<0.05, **p<0.01, ***p<0.001 SE: Standard error. SC: Superior colliculus. Display items in panels **d** and **e** were created using BioRender (BioRender.com/1m4j5bz).

Note: The functional relationships among pulvinar divisions during threat learning should be interpreted as computational dependencies derived from statistical associations. These effects may reflect indirect interactions mediated by corticothalamic and thalamocortical pathways (e.g., via visual cortex), rather than direct inter-nuclear connectivity. Elucidating the underlying anatomical mechanisms will require future studies.

These underlying dynamics and the activation timing in the trial-wise analyses (overview in **Fig. 3c** and details in **Fig. 1b**) led us to propose a possible functional model of the relationships between pulvinar divisions. Specifically, the increased activation induced by CS+ during trial 1 in the medial ($p_{FDR} = 0.026$), inferior ($p_{FDR} = 0.008$), and lateral pulvinar ($p_{FDR} = 0.01$) suggests that these divisions process CS+ information before the anterior pulvinar, which exhibited a delayed BOLD response starting in trial 2 only ($p_{FDR} = 0.44$ in trial 1, while $p_{FDR} = 0.000001$ in trial 2). The medial pulvinar mediates connections within the pulvinar network and is associated with elevated centrality values. Together, these findings highlight a possible hub role of the medial pulvinar, integrating CS+ information at a higher level, compared to sensory-driven processing in the inferior and lateral pulvinar. This possible hierarchical organization is supported by previous evidence showing that the inferior and lateral pulvinar are associated with processing basic sensory information (Berman and Wurtz, 2010, 2008; Cortes et al., 2024), while the medial pulvinar is implicated in higher-order processing, including attention and working memory (Homman-Ludiye and Bourne, 2019) (**Fig. 3d**).

This data-driven approach provides a possible new perspective on the pulvinar divisions' functional specialization during associative threat learning. We hypothesize that, during this process, the activation in the medial pulvinar mediates the functional relationships between the inferior and lateral divisions with the anterior pulvinar (**Fig. 3e**). To test this hypothesis, we conducted a mediation model analysis and evaluated

the model robustness using the k-fold cross-validation method. The sample (*N* = 293) was randomly divided into three groups of 91, 96, and 106 subjects. Each sub-sample was used to test the model while the remaining sub-samples were used to assess the stability of the parameter estimates. This resulted in six iterations of the mediation model. The 95% CI for each path coefficient were estimated using 10,000 bootstrapping replications. This choice was made to achieve greater precision and stability in the CI estimates. Finally, applying the mediation analysis to another independent cohort (*N* = 114) confirmed the model's applicability. The results in **Figs. 3f-3g** support the mediating role of the medial pulvinar, demonstrating significant indirect computational paths from the lateral and inferior divisions to the anterior pulvinar (*both p-values* < *0.05*). Detailed statistical parameters are provided in **Supplementary Tables 3**.

Pulvinar divisions and MD seem to support extinction learning while preserving threat memory across contextual cues

After our comprehensive examination of the thalamic contribution to threat conditioning, we moved forward to examine the neural contributions of the same regions to extinction learning. During this phase, at the block level, we found higher activation in response to CS+E than CS- across all pulvinar divisions ($t_{(319)} = 2.02–2.85$, $p = 0.005–0.0.044$, $d = 0.11–0.15$) and MD ($t_{(319)} = 3.04$, $p = 0.003$, $d = 0.17$). The trial-wise analysis indicated that these differences were primarily driven by the first two trials in the pulvinar; across pulvinar divisions, all second-trial comparisons survived FDR correction ($p_{FDR} < 0.05$, **Fig. 4b**), whereas the MD showed no significant effects at the trial-wise level (*all* $p_{FDR} > 0.05$, **Fig. 4c**). Thus, these nuclei—particularly the pulvinar—showed possible engagement in rapid extinction learning, whereas cumulative block-level differences in

the MD suggest a role in updating higher-order representations of the overall task structure. Detailed statistical parameters are provided in **Supplementary Tables 4-5**.

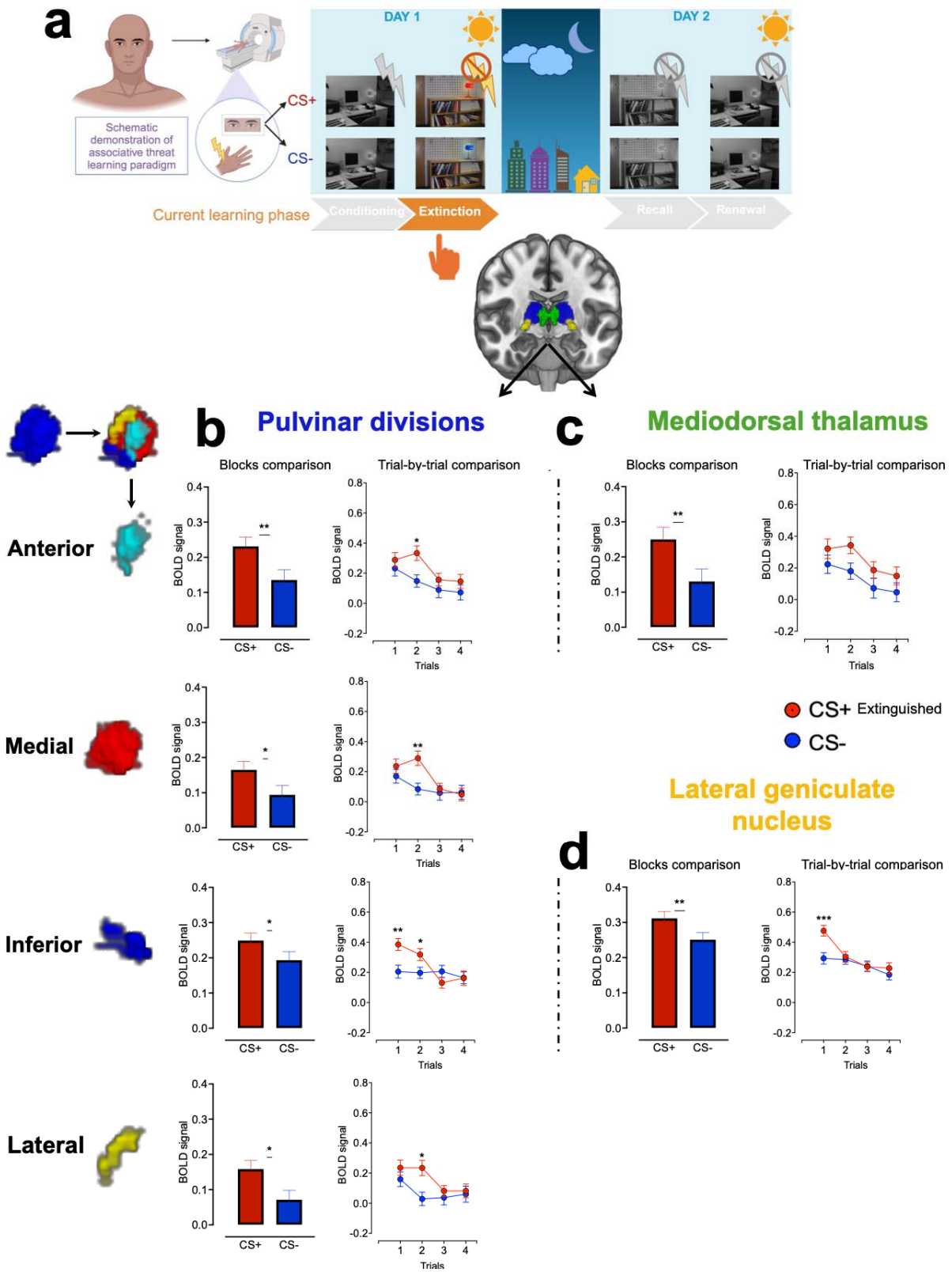

Fig. 4. Neural representation during extinction learning in pulvinar divisions, MD, and LGN. (a) Human extinction learning paradigm during fMRI.

(b–d) Means ± SE of activation in response to CS+ vs. CS− at both block-wise and trial-wise levels within pulvinar divisions (b), MD (c), and LGN (d). Block-level comparisons were assessed using paired *t*-tests, while trial-level effects were examined using a 2 × 2 repeated-measures ANOVA, followed by post hoc comparisons between CS+ and CS− across the four trials. Multiple comparisons were controlled using false discovery rate (FDR) correction. Extinction sample size: n = 320. Detailed statistical parameters are provided in **Supplementary Tables 4–5**.
*p < 0.05, **p < 0.01, ***p < 0.001. SE, standard error; MD, mediodorsal thalamus; LGN, lateral geniculate nucleus; Extinguished CS+, conditioned stimulus that no longer predicts shock; CS−, conditioned stimulus predicting no shock. Display items in panel (a) were created using BioRender (BioRender.com/iori2z1).

During extinction recall, the anterior pulvinar and MD exhibited similar functional patterns at both block and trial levels, with increased activation to CS+E compared to CS- during the first two trials (Anterior pulvinar: at the block level $t_{(411)} = 2.49$, $p = 0.013$, $d = 0.12$, $p_{FDR} = 0.026$ in trial 1 and $p_{FDR} = 0.026$ in trial 2; while $p_{FDR} > 0.05$ in trial 3 and 4. MD: at the block level $t_{(411)} = 4.3$, $p = 0.00002$, $d = 0.21$, $p_{FDR} = 0.000002$ in trial 1 and $p_{FDR} = 0.008$ in trial 2 while $p_{FDR} > 0.05$ in trial 3 and 4; **Fig. 5b-c**). In contrast, the medial and inferior pulvinar maintained the BOLD signal to CS+E in the second trial (both p_{FDR} in trial 2 < *0.01* while p_{FDR} in other trials > *0.05*). We observed no activation differences in the lateral pulvinar (*p-values* at the block and trial-wise level > *0.05*). See **Fig. 5b-c**, and Detailed statistical parameters in **Supplementary Tables 6-7**. These findings highlight the involvement of most pulvinar divisions and MD in sustaining threat memory under safe contextual cues and suggests association with retrieval suppression of the extinguished threat.

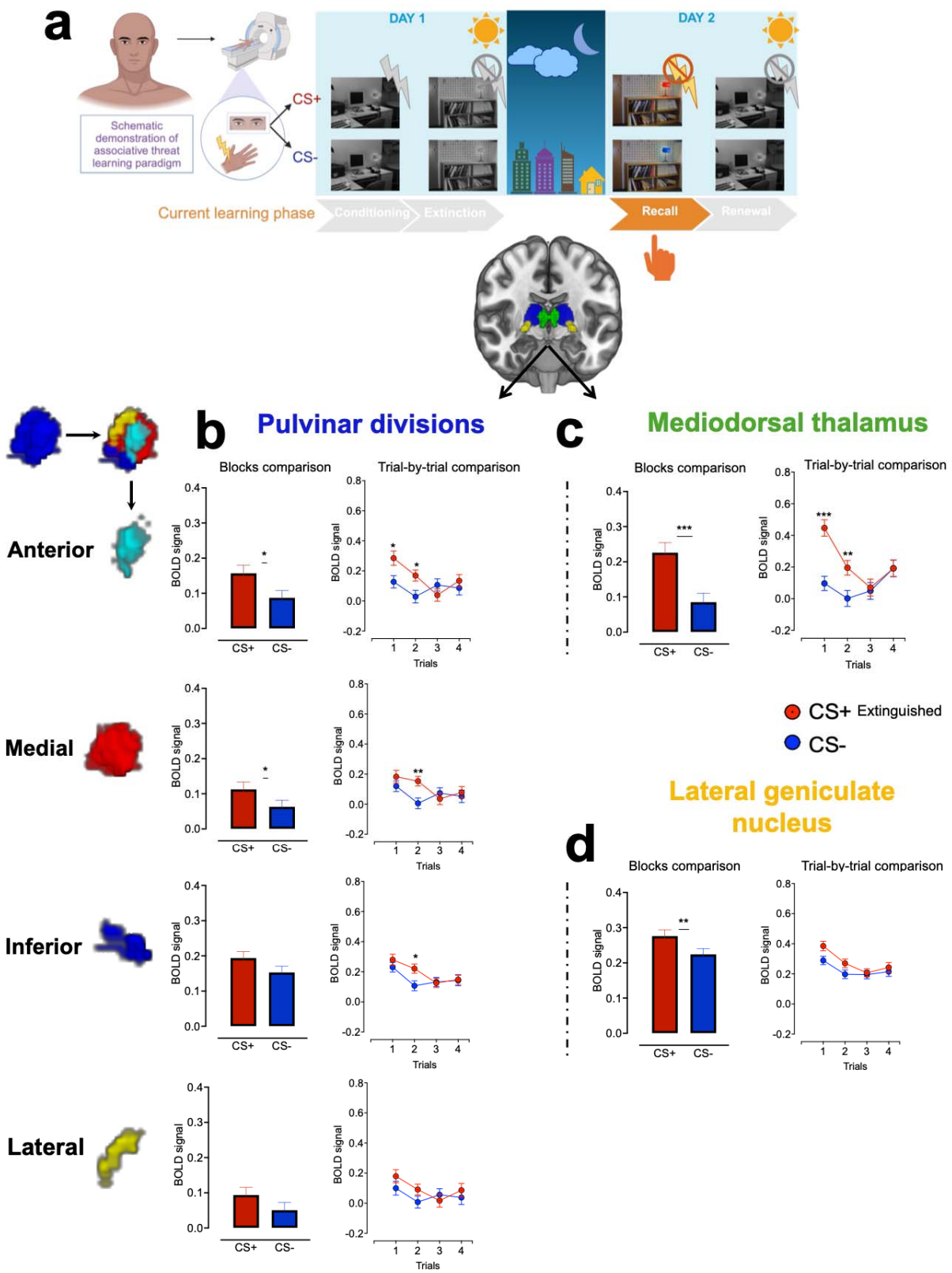

Fig. 5. Neural representation during extinction recall in pulvinar divisions, MD, and LGN. (a) Human extinction recall paradigm in the fMRI conducted within safe contextual cues.

(b–d) Means ± SE of activation in response to CS+ vs. CS− at both block-wise and trial-wise levels within pulvinar divisions (b), MD (c), and LGN (d). Block-level comparisons were assessed using paired *t*-tests, while trial-level effects were examined using a 2 × 2 repeated-measures ANOVA, followed by post hoc comparisons between CS+ and CS− across the four trials. Multiple comparisons were controlled using false discovery rate (FDR) correction. Extinction recall sample size: $n = 412$. Detailed statistical parameters are provided in **Supplementary Tables 6–7.**
*p < 0.05, **p < 0.01, ***p < 0.001. SE, standard error; MD, mediodorsal thalamus; LGN, lateral geniculate nucleus; Extinguished CS+, conditioned stimulus that no longer predicts shock; CS−, conditioned stimulus predicting no shock. Display items in panel (a) were created using BioRender (BioRender.com/imzbkmn).

Changing the contextual cues to the threat background in which the original associative learning occurred was associated with increased activation to CS+E on the first trial in the anterior pulvinar ($p_{FDR} = 0.0001$), lateral pulvinar ($p_{FDR} = 0.008$), and MD ($p_{FDR} = 0.0003$). No significant activation was observed in the medial or inferior pulvinar at the trial level (all $p_{FDR} > 0.05$ across trials; **Fig. 6b–c**; Detailed statistical parameters are provided in **Supplementary Tables 8-9**.

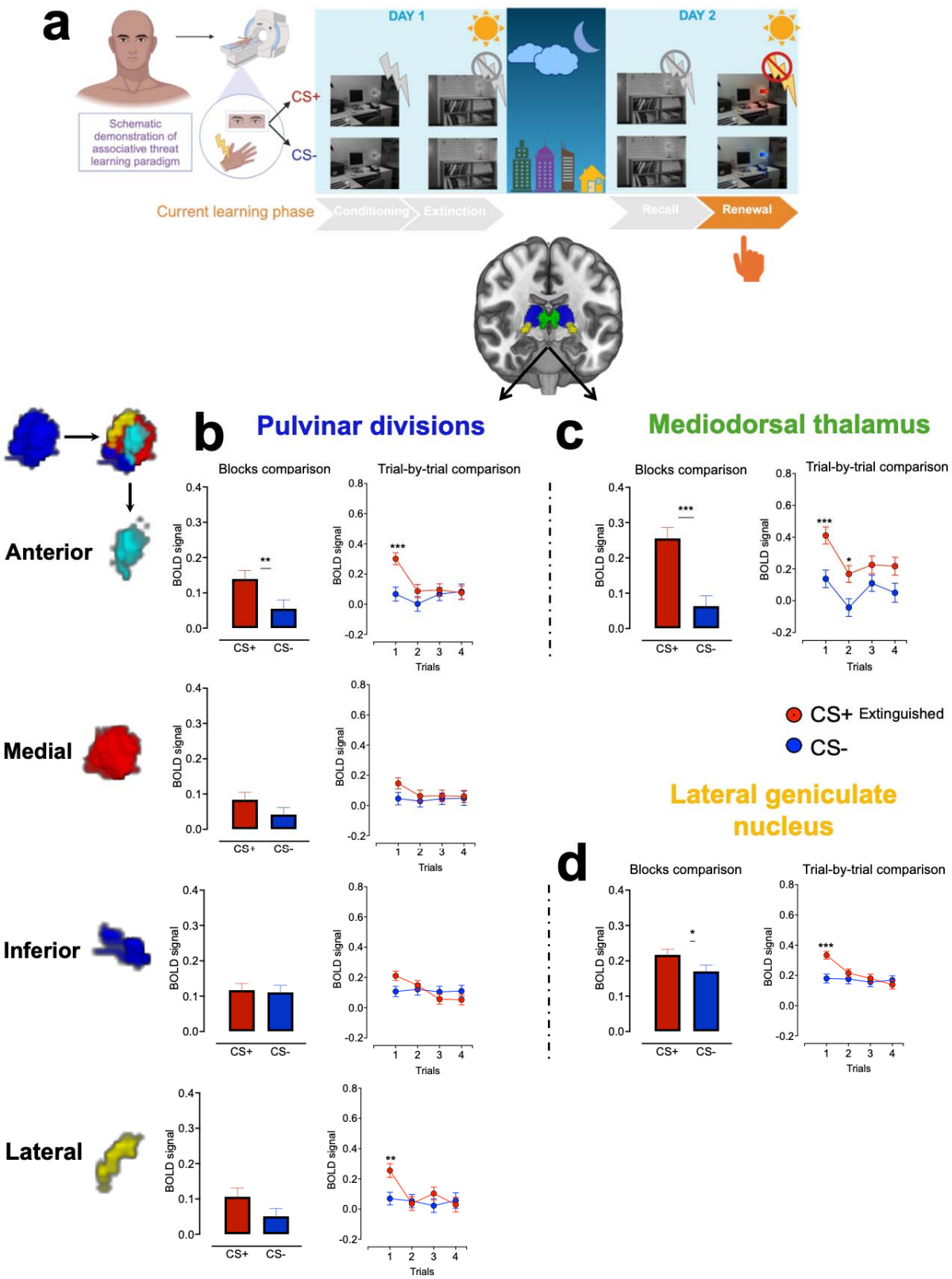

Fig. 6. Neural representation during threat renewal in pulvinar divisions, MD, and LGN. (a) Human threat renewal paradigm in the fMRI conducted within threat contextual cues in the original context where fear conditioning occurred.

(b–d) Means ± SE of activation in response to CS+ vs. CS− at both block-wise and trial-wise levels within pulvinar divisions (b), MD (c), and LGN (d). Block-level comparisons were assessed using paired *t*-tests, while trial-level effects were examined using a 2 × 2 repeated-measures ANOVA, followed by post hoc comparisons between CS+ and CS− across the four trials. Multiple comparisons were controlled using false discovery rate (FDR) correction. Threat renewal sample size: $n = 318$. Detailed statistical parameters are provided in **Supplementary Tables 8–9.**
*p < 0.05, **p < 0.01, ***p < 0.001. SE, standard error; MD, mediodorsal thalamus; LGN, lateral geniculate nucleus; Extinguished CS+, conditioned stimulus that no longer predicts shock; CS−, conditioned stimulus predicting no shock. Display items in panel (a) were created using BioRender (BioRender.com/r9bhqjs).

LGN activation patterns during threat learning are consistent with feedforward processing

At the block level, activation was greater for CS+ than CS− across conditioning ($t_{(292)} = 3.35$, $p = 0.0009$, $d = 0.19$), extinction ($t_{(319)} = 2.66$, $p = 0.008$, $d = 0.14$), recall ($t_{(411)} = 2.61$, $p = 0.009$, $d = 0.12$), and renewal ($t_{(317)} = 2.41$, $p = 0.016$, $d = 0.13$). These differences were primarily driven by the first trial ($p_{FDR} < 0.001$ in trial 1 across phases, except recall: $p_{FDR} = 0.056$), during which CS+ elicited increased BOLD responses that declined over subsequent trials. See **Figs. 1d**, **4d–6d**; Detailed statistical parameters are provided in **Supplementary Tables 2, 5, 7, and 9**. Collectively, these findings suggest that the LGN contributes to multiple phases of threat learning, with early trial responses consistent with feedforward visual processing.

Thalamic connectivity underlies threat learning and memory

We tested the relationships between each thalamic nucleus and core brain regions within the 'fear circuit' (Herry et al., 2010; Maren and Quirk, 2004; Milad et al., 2014; Milad and Quirk, 2012; Tovote et al., 2015) by seeding the nuclei to target the amygdala, hippocampus, ventromedial prefrontal cortex (vmPFC), subgenual anterior cingulate cortex (sgACC), and dorsal anterior cingulate cortex (dACC). During conditioning, the anterior pulvinar showed significant CS+ > CS− connectivity with the amygdala ($t_{(292)} = 2.46$, $p_{FDR} = 0.02$, $d = 0.14$), vmPFC ($t_{(292)} = 2.60$, $p_{FDR} = 0.02$, $d = 0.15$), and hippocampus ($t_{(292)} = 2.42$, $p_{FDR} = 0.02$, $d = 0.14$), but not with the dACC ($t_{(292)} = -0.95$, $p_{FDR} = 0.42$, $d = -0.05$) or sgACC ($t_{(292)} = 0.79$, $p_{FDR} = 0.42$, $d = 0.04$), **Fig. 7a**.

Additionally, to assess the robustness of the connectivity differences between CS+ and CS−, we performed nonparametric bootstrapping (10,000 resamples) and estimated the bias-corrected and accelerated (BCa) CI for the mean paired differences. The BCa intervals excluded zero for the amygdala (BCa 95% CI [0.002, 0.015]), vmPFC (BCa 95% CI [0.002, 0.017]), and hippocampus (BCa 95% CI [0.0016, 0.016]), but included zero for the dACC (BCa 95% CI [-0.01, 0.003]) and sgACC (BCa 95% CI [-0.005, 0.011]). The connectivity with the amygdala likely supports encoding the emotional valence of the newly learned CS-US associations, while the engagement of the hippocampus suggests prioritizing the contextual CS+ information. The connectivity with the vmPFC may underlie a process of top-down control release to enhance fear

expression or encode threat information for tracking and supporting decision-making in future encounters.

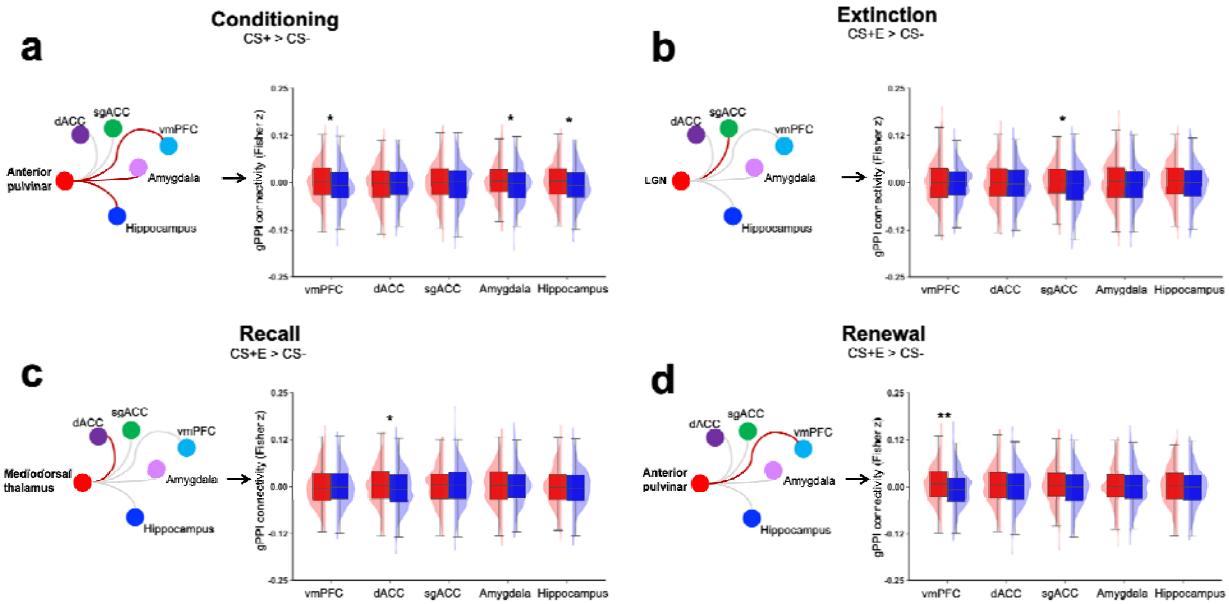

Fig. 7. Task-related thalamic connectivity with threat-related brain regions.
(a–d) Thalamic seed-to-ROI connectivity with the vmPFC, sgACC, dACC, amygdala, and hippocampus during conditioning (*n = 293*), extinction (*n = 320*), recall (*n = 412*), and renewal (*n = 318*). Red lines indicate significant positive connectivity for CS+ vs. CS− following FDR correction across regions ($p_{FDR} < 0.05$); gray lines indicate non-significant effects. For each panel, boxplots and kernel density estimates show the distribution of connectivity values for CS+ and CS−.
All thalamic regions of interest—including pulvinar divisions (anterior, inferior, lateral, medial), LGN, and MD—were analyzed using identical statistical procedures. To maintain figure clarity and focus on effects relevant to the main conclusions, we display only seeds showing at least one significant CS+ vs. CS− differences after FDR correction. Non-significant seeds did not show systematic condition-related differences.
*p < 0.05, **p < 0.01. MD, mediodorsal thalamus; LGN, lateral geniculate nucleus; vmPFC, ventromedial prefrontal cortex; sgACC, subgenual anterior cingulate cortex; dACC, dorsal anterior cingulate cortex; CS+, conditioned stimulus predicting shock; CS−, conditioned stimulus predicting no shock. gPPI; generalized psychophysiological interaction; connectivity values are Fisher z–transformed correlation coefficients.

During extinction learning, the LGN showed increased CS+E > CS− connectivity with the sgACC ($t_{(319)} = 2.60$, $p_{FDR} = 0.04$, $d = 0.14$). No significant differences were observed with other cortical regions, including the vmPFC ($t_{(319)} = 1.13$, $p_{FDR} = 0.32$, $d = 0.06$) and dACC ($t_{(319)} = 0.41$, $p_{FDR} = 0.68$, $d = 0.02$), or with limbic regions, including the amygdala ($t_{(319)} = 1.19$, $p_{FDR} = 0.32$, $d = 0.06$) and hippocampus ($t_{(319)} = 1.34$, p_{FDR} =

0.32, d = 0.07), **Fig. 7b**. Additionally, to assess the robustness of the connectivity differences between CS+E and CS−, we performed nonparametric bootstrapping analysis (10,000 resamples). The BCa interval for the mean differences excluded zero for the LGN–sgACC connectivity (BCa 95% CI [0.002, 0.017]), but included zero for all other regions: vmPFC (BCa 95% CI [−0.003, 0.011]), dACC (BCa 95% CI [−0.005, 0.009]), amygdala (BCa 95% CI [−0.002, 0.012]), and hippocampus (BCa 95% CI [−0.002, 0.012]).

Interestingly, in extinction recall and threat renewal, we found that the learners, i.e., the anterior pulvinar and MD seems to support either safe or threatening memory, depending on contextual cues. Specifically, during extinction recall, which occurs in the safe context, we observed increased connectivity to CS+E > CS- between the MD and dACC ($t_{(411)} = 2.60$, $p_{FDR} = 0.04$, $d = 0.12$) but not with other regions: amygdala ($t_{(411)} = 0.038$, $p_{FDR} = 0.97$, $d = 0.002$), hippocampus ($t_{(411)} = 1.2$, $p_{FDR} = 0.57$, $d = 0.059$), sgACC ($t_{(411)} = 0.18$, $p_{FDR} = 0.97$, $d = 0.009$), and vmPFC ($t_{(411)} = -0.49$, $p_{FDR} = 0.97$, $d = -0.02$), **Fig. 7c**.

These findings were robust to nonparametric bootstrapping (10,000 resamples): the BCa 95% confidence interval for the mean connectivity difference excluded zero for the MD–dACC (BCa 95% CI [0.002, 0.015]), but included zero for the amygdala (BCa 95% CI [−0.006, 0.006]), hippocampus (BCa 95% CI [−0.002, 0.010]), sgACC (BCa 95% CI [−0.006, 0.007]), and vmPFC (BCa 95% CI [−0.008, 0.005]).

During threat renewal, we observed increased CS+E > CS− connectivity between the anterior pulvinar and the vmPFC ($t_{(317)} = 3.47$, $p_{FDR} = 0.002$, $d = 0.19$). In contrast, no

significant CS+E > CS− connectivity differences were observed between the anterior pulvinar and other regions, including the amygdala ($t_{(317)} = 0.59$, $p_{FDR} = 0.55$, $d = 0.03$), hippocampus ($t_{(317)} = 1.09$, $p_{FDR} = 0.44$, $d = 0.06$), sgACC ($t_{(317)} = 1.90$, $p_{FDR} = 0.14$, $d = 0.10$), and dACC ($t_{(317)} = 0.92$, $p_{FDR} = 0.44$, $d = 0.05$), **Fig. 7d**.

Nonparametric bootstrapping (10,000 resamples) supported the robustness of these results. The BCa 95% confidence interval for the mean difference excluded zero only for the anterior pulvinar –vmPFC connectivity (BCa 95% CI [0.005, 0.02]), whereas intervals for the amygdala (BCa 95% CI [-0.005, 0.01]), hippocampus (BCa 95% CI [-0.003, 0.012]), sgACC (BCa 95% CI [-0.0002, 0.015]), and dACC (BCa 95% CI [-0.003, 0.01]) all included zero.

Discussion

We examined the neural representation of associative threat learning within the pulvinar divisions, LGN, and MD, providing new insights into thalamic involvement in this adaptive behavior in humans. We identified distinct thalamic activation profiles during threat learning and memory. The anterior pulvinar and MD exhibited parallel activation patterns consistent with associative learning, raising a possibility that these nuclei are involved in automatic survival responses and deliberate threat processing. We propose a novel hierarchical computational model for processing threat information in the pulvinar divisions. The medial pulvinar seems to mediate basic sensory information from the inferior and lateral divisions to the anterior pulvinar for higher-order integrative learning. Both pulvinar and MD were involved in extinction and showed activation consistent with the salience processing of threat-related memories during extinction

recall and threat renewal. The LGN primarily seems to represent feedforward processing, anticipating upcoming visual stimuli throughout threat learning. We integrated these insights into schematic models underlying the emotional and behavioral expression of threat learning and memory, providing a possible neural framework for studying related human behaviors (**Fig. 8a-b**).

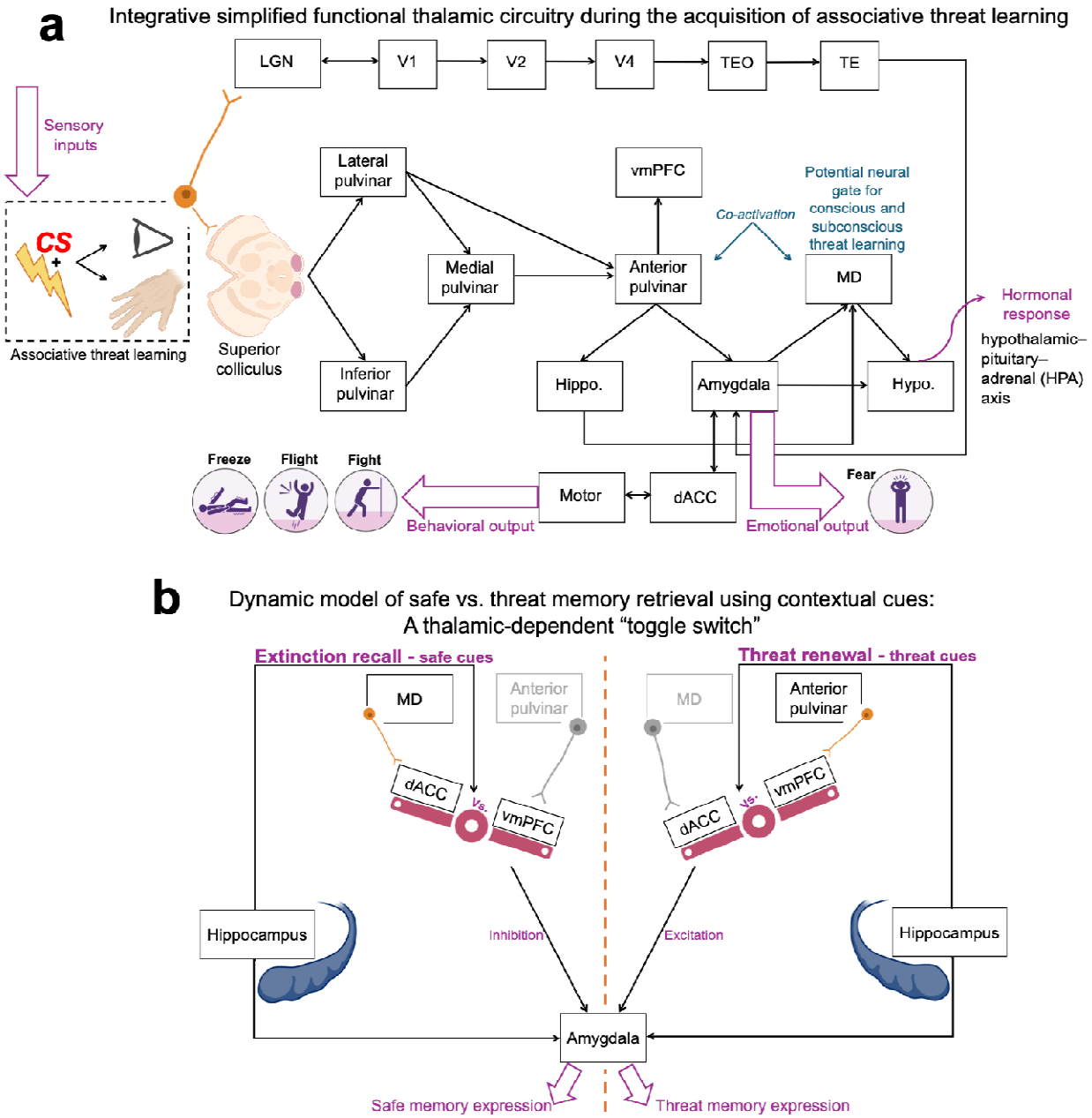

Fig 8. Neurobehavioral models of thalamic involvement in associative threat learning and memory.

a. Schematic illustration of thalamic circuitry during the acquisition of associative threat learning, highlighting interactions between pulvinar divisions, MD, and LGN with key brain regions involved in fear expression.

b. A thalamic-dependent "toggle switch" regulates the retrieval of safety vs. threat-related memory. The MD-dACC connectivity modulates the interaction between dACC and vmPFC, promoting vmPFC dominance during extinction recall. In contrast, the anterior pulvinar-vmPFC connectivity promotes dACC dominance, enhancing the expression of threat memory during threat renewal.

MD: Mediodorsal thalamus. LGN: Lateral geniculate nucleus. vmPFC = Ventromedial prefrontal cortex. dACC = Dorsal anterior cingulate cortex. V1, V2, and V4: Primary, secondary, and fourth visual areas. TEO, TE: Temporal cortex regions. Hypo.: Hypothalamus. Hippo.: Hippocampus.

Created using BioRender (BioRender.com/5pdf54t).

Note (panel a): Known pulvinar–cortical connections, as well as sensory input pathways (e.g., visual inputs via the retina/LGN and nociceptive inputs via the spinothalamic tract), are not explicitly shown. These connections are well established anatomically but were omitted due to their heterogeneity and incomplete characterization at the level of pulvinar subnuclei. Their absence should not be interpreted as a lack of anatomical or functional relevance.

The anterior pulvinar and MD co-activation patterns may demonstrate a parallel contribution to threat learning. The similar activation in response to CS+ and CS- at the first trial, followed by a heightened activation specific to CS+ in the next trial, is likely consistent with the acquisition of the CS-US association. The responses' similarity to both types of CS during the first trial suggests a possible equivalent initial emotional valence. The gradual increase in the activation, specifically to CS+ by the end of trial 1, suggests a shift in the emotional valence of CS+, likely indicating rapid associative learning. This suggested parallel specialized role highlights the anterior pulvinar and MD as central thalamic hubs for integrating CS-US information, apparently specifying the two-system model proposed by LeDoux and Pine (LeDoux and Pine, 2016). Briefly, fear processing, according to this model, involves two distinct pathways: a subcortical route for rapid threat detection and species-specific defense reactions and a cortical route for deliberate threat evaluation and conscious fear experience. Learning within the anterior pulvinar appears to support rapid automatic processes of the subcortical pathway. This aligns with previous reports that demonstrated the role of the pulvinar-amygdala pathway in encoding negative emotions in humans (Bridge et al., 2015; Koller et al., 2019; Kragel et al., 2021; McFadyen et al., 2019; Rafal et al., 2015). Indeed, the increased connectivity that we observed with the amygdala, hippocampus, and vmPFC during threat conditioning aligns with this role. Specifically, in this context, the amygdala likely contributes to tagging a negative emotional valence to CS+ and triggering fight, flight, or freeze reactions (Adolphs et al., 1994; Costa et al., 2022; Fanselow and

LeDoux, 1999; Phelps and LeDoux, 2005; Wen et al., 2024, 2022). The connectivity with the hippocampus likely facilitates contextual encoding of the environmental characteristics of the aversive event (Maren, 2001; Maren et al., 2013). In turn, the connectivity with vmPFC appears to support salience encoding for tracking future encounters with the learned threat (Battaglia et al., 2020) and facilitating decision-making and emotion regulation (Milad and Quirk, 2012, 2002; Nejati et al., 2021).

On the other hand, MD activation aligns more closely with the cortical pathway, facilitating conscious and deliberate threat encoding. This is supported by our findings, which showed increased MD activation in response to CS+, compared to the anterior pulvinar during threat learning, probably indicating a different level of threat processing. Additionally, this interpretation is supported by substantial evidence pointing to well-established anatomical pathways connecting the MD with the PFC and demonstrating its role in decision-making and learning (Behrens et al., 2003; Hwang et al., 2020; Li et al., 2022; Wolff and Halassa, 2024). This is along with reports that underscored the contribution of the MD-cortical loops to supporting a conscious experience in humans (Griffiths et al., 2022; Whyte et al., 2024). Together, these findings align with the framework of the two-system model of threat processing, highlighting the anterior pulvinar and MD possible role in the acquisition of the CS-US association and suggesting that these regions might be involved in a parallel conscious vs. unconscious learning proposed by LeDoux and colleagues (LeDoux and Pine, 2016), **Fig. 8a**.

This suggested rapid initial acquisition of the predictive value of the CS+ is thought to be pronounced during the first two trials. The attenuated CS+ vs. CS− differentiation on the

third trial specifically in the pulvinar may reflect a decreased requirement for differential thalamic engagement once the initial association has been acquired, or an initial survival fear reaction is expressed. Notably, because the MD sustained the BOLD response to the CS+ in the third trial which may indicate involvement of this nucleus in the consolidation or stabilization of the learned association. This aligns with the well-established MD-PFC circuit involved in cognitive processes (Wolff and Halassa, 2024). Additionally, in a previous study using a similar paradigm, we observed sustained CS+ vs. CS− differentiation on the third trial in the nucleus reuniens, as well (Tuna et al., 2025). These findings suggest that trial-dependent learning dynamics may vary across thalamic nuclei rather than reflecting a uniform thalamic learning signal. Together, while our paradigm does not inherently distinguish between different stages of learning, such as early acquisition and stabilization, our findings are consistent with stronger associative learning–related engagement during the first two trials, with a reduced differential response by the third trial that may reflect the involvement of different neural processes.

We suggest that different pulvinar divisions seems to hierarchically contribute to the integration of CS-US association in the anterior pulvinar, likely through bottom-up processing. Although the anatomical projections within pulvinar divisions are not well characterized, our functional data-driven approach suggests a possible critical role for the medial pulvinar as a hub region facilitating communication among pulvinar divisions. Specifically, our trial-wise analysis suggested that the medial, inferior, and lateral pulvinar process threat-related information earlier than the anterior pulvinar. The anterior pulvinar exhibited activation in response to CS+ starting at trial 2, whereas

activation in the other pulvinar divisions occurred as early as trial 1. Network modeling further identified the medial pulvinar as a potential central hub, likely facilitating communication with other divisions. Using a robust mediation model, we demonstrated that the medial pulvinar appears to mediate the relay of sensory CS+ information from the inferior and lateral pulvinar to the anterior division probably for higher-order integrative learning. This hierarchical model aligns with previous evidence suggesting that the inferior and lateral pulvinar are associated with processing basic sensory information (Berman and Wurtz, 2010, 2008; Bridge et al., 2015; Cortes et al., 2024). While the medial pulvinar, which receives projections from deep layers of the superior colliculus, supports more advanced functions such as attention and working memory (Bridge et al., 2015; Homman-Ludiye and Bourne, 2019). These findings seem to be indicative of novel insights into the functional specialization of pulvinar divisions during threat learning, suggesting feedforward functions in the inferior and lateral pulvinar and high-order integration in the anterior pulvinar.

Yet, these intrapulvinar relationships should be understood as a functional and computational model, reflecting statistical dependencies among pulvinar divisions during threat learning, rather than as evidence of direct monosynaptic anatomical connections. Because detailed inter-nuclear anatomical connectivity within the pulvinar remains incompletely characterized, our analysis does not presuppose strong direct excitatory projections between subnuclei. Instead, our findings are intended to highlight candidate functional relationships within the pulvinar during conditioning, rather than to provide a definitive anatomical map.

The pulvinar divisions and the MD showed patterns consistent with engagement during extinction learning. In the pulvinar, responses to the CS+E remained relatively stable across the first two trials and subsequently attenuated; a pattern that may reflect rapid updating of threat value. In contrast, MD activation patterns were more consistent with representing the broader task structure across blocks. Despite the involvement in extinction, these nuclei seem to remain sensitive to the extinguished threat during safe contextual cues in extinction recall and exhibited increased activation to CS+E when the stimuli were presented within threat contextual cues during renewal. This pattern suggests engagement in retrieval suppression during recall, along with threat salience processing during both recall and renewal. The MD and pulvinar is suggestive to continue to monitor and evaluate the extinguished threat across contexts rather than relying on static safety memories. These nuclei appear to engage in dynamic evaluation, learning, and decision-making during future encounters with the extinguished CS+.

The classical Pavlovian model proposes that extinction learning forms a new safety memory, competing with the original threat memory acquired during conditioning (Bouton, 2002; Milad and Quirk, 2002). Contextual changes often favor either the safe or threat memory. Maren and colleagues (Maren et al., 2013; Maren and Quirk, 2004) described the hippocampal-prefrontal-amygdala model for contextual memory. The hippocampus projects to the basolateral amygdala, vmPFC, and dACC. Although the vmPFC supports safe memory recall by projecting to the intercalated cells (which inhibit the central nucleus of the amygdala), the dACC enhances threat memory renewal by projecting to the basolateral amygdala (which activates the central nucleus of this

structure). Our findings suggest that the thalamic connectivity may influence the balance between recalling safe vs. threat memory by modulating the competitive interaction between the dACC and vmPFC. The increased MD-dACC connectivity during recall may reflect a thalamic-driven reconfiguration of prefrontal circuits, enabling the vmPFC to become more functionally dominant and promote safe memory expression. On the other hand, the anterior pulvinar-vmPFC connectivity during renewal may destabilize the influence of the vmPFC, allowing the dACC to regain dominance and facilitate threat memory retrieval. The thalamic connectivity, thus, appears to orchestrate a context-dependent functional balance between the dACC and vmPFC. This balance may serve as a neural "toggle switch" between safe vs. threat memory expression. We integrated this suggested flexible responding mechanism to changing environmental contexts in Maren's circuit model of emotional memory, **Fig. 8b**.

The LGN contribution seems to be consistent across most phases of threat learning, as evidenced by our trial-wise results, which pointed to increased activation in response to CS+ that diminished immediately after the first trial. This distinct pattern might indicate readiness for imminent visual stimuli, regardless of the actual emotional valence, as conditioning, extinction, and renewal elicited similar BOLD patterns. This aligns with previous studies that highlighted the LGN's engagement in selective attention and anticipation of visual stimuli (Mahoney and Schmidt, 2024; O'Connor et al., 2002; Saalmann and Kastner, 2009). The LGN, as a first-order nucleus (Cortes et al., 2024), primarily supports a feedforward function, while the pulvinar, as part of the visual thalamus, is more closely associated with a broader functional processing (Cortes et al.,

2024) including emotional valences of stimuli. This distinction likely underscores their joint but specialized contributions to adaptive threat responses.

Limitations and Future Directions

Although distinct thalamic roles in threat learning have been proposed, fMRI data do not fully capture the complexity of this structure. Pulvinar divisions, MD, and LGN each contain diverse neuron subtypes and finer anatomical divisions that may serve distinct functions. Importantly, the absence of CS+ vs. CS− differences in BOLD activation should not be interpreted as a lack of stimulus-specific processing, as such distinctions may occur without changes in overall activation detectable by fMRI. Future advancements, including higher-resolution human brain atlases, may improve anatomical and functional precision. Additionally, since different sensory modalities preferentially engage distinct thalamic nuclei, the specific thalamic roles we investigated may not be consistent across different experimental designs, particularly in studies using auditory threat learning. The observed relationships among pulvinar divisions during conditioning are purely functional and do not distinguish direct inter-nuclear interactions from indirect coupling mediated by corticothalamic and thalamocortical pathways, including visual cortical regions. Thus, the pulvinar model may reflect indirect cortical loops, or currently undocumented inter-nuclear interactions, or a combination of both. Finally, given the indirect nature of fMRI data and the absence of direct brain signal manipulations, our findings should not be interpreted as evidence of causality. Further research is needed to examine causal mechanisms underlying dynamic neural representations of threat learning within the thalamus.

This study's insights raise critical future questions at the intersection between neuroscience, psychology, and mental health, particularly regarding the neural mechanisms underlying associative threat learning. First, the crucial role of the anterior pulvinar and MD during the acquisition of the CS-US association sets the stage for developing more precise brain interventions, focusing on refining maladaptive fear reactions. Future studies could explore the optimal parameters for applying non-invasive techniques to inhibit these thalamic regions during or immediately after fear conditioning. This approach is particularly promising for clinical populations and individuals exposed to high-risk environments, such as emergency medical staff, firefighters, and paramedics.

Second, the distinct contributions of the MD-dACC and anterior pulvinar-vmPFC circuits to safe vs. threatening memory suggest potential intervention targets for prioritizing one memory over the other. Activating the MD-dACC may facilitate recalling safe memories while engaging the anterior pulvinar-vmPFC circuits might reinforce fear memory pathways. Experimental designs in controlled laboratory settings could target these circuits to identify specific parameters for enhancing safe memories. This avenue holds therapeutic potential, particularly in conditions such as aviophobia; activating a safe memory circuit before a flight might suppress fear relapse during the actual flight.

Third, the anterior pulvinar-MD relationships raise questions related to the wide range of human behaviors, focusing on understanding neural gateways between conscious and unconscious learning. Studying the dynamics of information flow across this promising pathway could deepen our understanding of how sensory and cognitive information

transits between conscious and unconscious states and brings about behaviors. This research line may pave the way for innovative learning methods and reveal neurocognitive mechanisms underlying the acquisition of new information in humans.

Together, our findings and emerging research directions underscore the thalamic nuclei's vital role as a hub for fear acquisition and memory processing, offering possible avenues for theoretical advancements and clinical applications. As the primary neural gateway to the human brain, the thalamus is the first station for all sensory information except olfactory inputs. Modulating its functions is potentially an impactful strategy to induce widespread functional changes across the brain and influence different mental and behavioral expressions.

Materials and Methods

Participants

We analyzed fMRI data of 293 participants during fear conditioning, 320 during extinction, 412 during extinction recall, and 312 during threat renewal. Participants included individuals of both sexes, aged 18–70 years (mean ± SD = 32.17 ± 13.1 years). All participants were proficient in English, right-handed, and had normal or corrected-to-normal vision. The exclusion criteria included a history of seizures or significant head trauma, current substance abuse or dependence, metal implants, pregnancy, breastfeeding, or positive urine toxicology screen for drugs of abuse. We followed the latest version of Helsinki's declaration, and all procedures were approved by the Partners HealthCare Institute Review Board of the Massachusetts General Hospital, Harvard Medical School. All subjects provided written informed consent before

taking part in the study. Results from this dataset have been published elsewhere with a different focus (Marin et al., 2020; Wen et al., 2024, 2022). The current results are novel and have not been previously published.

Experimental procedure

Threat learning: Participants underwent two-day sessions of a validated threat learning paradigm in the MRI scanner (**Figs. 1a, 4a-6a**); the experimental contexts consisted of visual scenes on a computer display. On day 1, participants underwent a Pavlovian conditioning phase in which a neutral stimulus in context A (e.g., red light in an office) was paired with a 500ms electric shock (US) with a partial reinforcement rate of 62.5%. Another neutral stimulus in context A (e.g., blue light in the same office) was also presented but never paired with the shock (CS-). On the same day, we also conducted extinction learning in context B (e.g., red and blue lights in a casebook background). The stimuli were repeatedly presented with the removal of the expected reinforcement (i.e., no shock).

On day 2, the participants underwent two phases of a memory test. The first is extinction recall, including presenting the extinguished CS+ and CS- within safe contextual cues (context B used during extinction learning). The second is threat renewal, in which the stimuli were presented within threat contextual cues (context A used in conditioning). Both extinction recall and threat renewal included no US reinforcement. Across phases of threat learning, the duration of each trial was 6s, and the inter-trial intervals with a fixation screen ranged between 12 to 18 s (15 s on average). Finally, to control for potential confounds related to the experimental design,

we pseudorandomized and counterbalanced the order of CS+ and CS− across phases and between subjects.

Shock level: The shock intensity used in the fear learning paradigm was determined during a pre-experiment calibration. Electrodes were attached to the participant's right hand, and stimulation began at a low level (0.1 mA), gradually increasing in small increments. After each increment, participants verbally rated their discomfort. The procedure continued until the participant identified a level they described as "highly annoying but not painful." This individualized intensity was then used for that participant throughout the experiment. For safety and ethical reasons, the maximum intensity was capped at 20 mA, and no participant received a shock above this limit.

Instructions to the participants: Each visual stimulus in our paradigm was first shown to participants for 6 seconds. This initial presentation served as habituation, allowing us to isolate the responses to genuinely new stimuli. Before the experiment began, participants were informed that they would see pictures illuminated with different colored lights, such as red or blue. During the experiment, some pictures might be paired with an electric shock, while others might not. Participants were instructed to pay attention to whether a specific color or pattern was associated with the shock. These instructions were adopted from previous studies in which our group developed this paradigm and found them highly effective for human learning. We therefore used the same approach in the current experiment. These instructions were provided throughout all phases of threat learning, and participants were informed that any shocks delivered would be at the same intensity determined on Day 1.

MRI data acquisition and preprocessing

Two MRI settings were used to acquire the neuroimaging data. The first is a Trio 3T whole-body MRI scanner (Siemens Medical Systems, Iselin, New Jersey) using an 8-channel head coil. The functional data in this setting were acquired using a T2* weighted echo-planar pulse sequence with these parameters: TR = 3.0s, TE = 30 ms, slice number = 45, voxel size = 3 × 3 × 3 mm^3. The second setting was also in the same scanner using a 32-channel head coil. The functional images were obtained using a T2* weighted echo-planar pulse sequence using TR = 2.56s TE = 30 ms, slice number = 48, voxel size = 3 × 3 × 3 mm^3. The anatomical brain images were collected using a T1-weighted MP-RAGE pulse sequence, parcellated into 1 × 1 × 1 mm^3 voxels. Elastic bands were affixed to the head coil device to reduce head motions.

Using the default pipeline in fMRIPrep, version 20.0.2 (Esteban et al., 2020, 2019), we preprocessed the data and applied correction of slice timing, realignment of the functional images, and coregistration. In addition, spatial normalization was performed with data normalized to Montreal Neurological Institute (MNI) space and resampled to a 2 × 2 × 2 mm^3 voxel grid, followed by spatial smoothing with a 6-mm full-width at half-maximum Gaussian kernel**.**

Activation analyses

We applied the least-squares-based generalized linear model (GLM) for each participant to estimate the BOLD response to CS+ and CS- using SPM 12. We estimated the beta values for each voxel during each learning phase in the paradigm. Overall, the model included 32 regressors for the CS+ and CS-, a regressor for the context, and a regressor for shock in conditioning but not in other phases. The GLM

also included six head movement parameters (x, y, z directions, and rotations). This first-level analysis resulted in contrast maps that we used to estimate the variability of these maps across all subjects at the group-level analysis. We then used the contrast maps from the group-level analysis to extract the averaged values across the voxels within predefined masks of the pulvinar divisions, MD, and LGN. These outputs were used to compare the BOLD response during the first four CS+ and four CS- trials. We averaged the activation for each CS type across trials to obtain the block-level activation.

Statistical tests

All statistical analyses were conducted using JASP versions 0.18.3 and 0.19.3 (JASP Team, 2024).

Activation differences across threat-learning phases

For each threat-learning phase, block-level activation was defined as the average activation across trials. Differences in activation between CS+ and CS− conditions within each thalamic nucleus were assessed using paired-samples t-tests.

To examine trial-wise activation dynamics, we conducted 2 × 2 RM-ANOVAs with condition (CS+, CS−) and Trial as within-subject factors. Assumptions of sphericity were evaluated using Mauchly's test, and Greenhouse–Geisser corrections were applied when violations were detected. Significant effects were followed by post hoc comparisons at the trial level, with false discovery rate (FDR) correction applied to control for multiple comparisons.

Network analyses of within-pulvinar relationships during conditioning

Network analyses examined functional relationships between pulvinar divisions. Nodes corresponded to block-level CS+ minus CS− activation estimates for each pulvinar division, yielding four nodes (one per division). Networks were estimated using a Gaussian graphical model with EBICglasso (LASSO regularization) based on Pearson correlation matrices, with the EBIC tuning parameter set to $\gamma = 0.5$. Edge weights represent partial correlations.

Three centrality measures were computed on the estimated weighted partial-correlation network: node strength, defined as the sum of absolute edge weights connected to a node; closeness, defined as the inverse of the average shortest path length from a node to all other nodes; and betweenness, defined as the proportion of shortest paths between all node pairs that pass through a given node. Shortest paths were computed using inverse edge weights. Centrality indices were normalized. Network accuracy and centrality stability were assessed using nonparametric bootstrapping (10,000 iterations) to estimate CI for edge weights and centrality measures. All network analyses were conducted using the default settings unless otherwise specified, following Epskamp et al., (2018)

Mediation analyses of within-pulvinar relationships during conditioning

Mediation models were estimated using the lavaan package (Rosseel, 2012) with maximum likelihood estimation. All variables were z-standardized prior to analysis. Block-level activation estimates from the inferior and lateral pulvinar were entered as

predictors, activation in the medial pulvinar as the mediator, and activation in the anterior pulvinar as the outcome.

To assess the robustness and generalizability of the mediation effects, we conducted K-fold cross-validation ($K = 3$). The full sample ($N = 293$) was randomly partitioned into three non-overlapping sub-samples ($n = 91, 96,$ and 106). In each iteration, the mediation model was estimated in one sub-sample, and consistency of indirect effects was evaluated across the remaining sub-samples, yielding six cross-validation iterations.

Indirect effects were evaluated using BCa bootstrap CI based on 10,000 resamples, as recommended by Biesanz et al., (2010). An indirect effect was considered significant when the 95% confidence interval did not include zero and $p < 0.05$.

Connectivity analyses

Functional connectivity was computed using the CONN toolbox (version 22.a) for MATLAB (Nieto-Castanon and Whitfield-Gabrieli, 2022; Whitfield-Gabrieli and Nieto-Castanon, 2012). Anatomical images were segmented into gray matter, white matter, and CSF, and functional images were preprocessed following CONN's default pipeline, including realignment, slice-timing correction, and coregistration. Functional images were then denoised using a component-based noise correction approach, along with regression of motion parameters and outlier scans, followed by band-pass filtering of the BOLD time series between 0.008 and 0.09 Hz.

Following standard denoising procedures—including regression of motion- and physiology-related confounds and temporal filtering—condition-dependent connectivity effects were inferred from subject-level generalized psychophysiological interaction (gPPI) contrast estimates rather than from raw time-series correlations. This GLM-based framework reduces the likelihood that observed PPI effects reflect differences in temporal autocorrelation or spectral properties across regions rather than genuine task-dependent interactions.

For ROI-to-ROI analyses, the pulvinar divisions, MD, and LGN were defined as seed regions, and connectivity was assessed with target regions including the dACC, sgACC, vmPFC, amygdala, and hippocampus. At the first level, functional connectivity between ROIs was computed as Pearson correlation coefficients of the denoised BOLD time series, which were then Fisher z-transformed to normalize the distribution. To estimate task-related connectivity, the first eight trials of each condition were analyzed as a block.

At the second level, these Fisher z-transformed connectivity values were entered into group-level GLMs to evaluate condition-dependent differences. Paired t-tests were used to assess differences between conditions, with false discovery rate (FDR, $p < 0.05$) correction for multiple comparisons. To provide robust estimates of mean differences, 10,000 bootstrap resamples were used to calculate the BCa CI.

Masks

We defined the pulvinar divisions, MD, and LGN nuclei using predefined masks based on the neuroanatomical guidelines of the Automated Anatomical Labelling Atlas(Rolls et al., 2020). We also applied the same atlas guidelines to determine the masks for the two other anatomical regions; the amygdala and hippocampus. The masks for the functional regions were created using Neurosynth (Yarkoni et al., 2011) and the keyword ‘conditioning.’ For each region, we created 8mm spheres around the following identified peak coordinates: vmPFC (MNIxyz = −2, 46, −10), sgACC (MNIxyz = 0, 26, −12), and dACC (MNIxyz = 0, 14, 28).

Acknowledgments

This work was supported by the National Institute of Mental Health grants R01MH123736, R01MH125198, R33MH111907, R01MH097880, and R01MH097964 to M.R.M.

Data and Code Availability

The data supporting the findings of this study are part of a larger, actively used database and cannot be publicly released at this time due to ongoing projects. The data underlying the analyses reported in this manuscript are available from the corresponding author upon reasonable request and subject to standard data-use agreements.

All scripts used for preprocessing and analysis, including workflows implemented in SPM12, and CONN, are available upon request to facilitate replication and reuse.

References

Adolphs R, Tranel D, Damasio H, Damasio A. 1994. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature 1994 372:6507 **372**:669–672. DOI: https://doi.org/10.1038/372669a0, PMID: 7990957

Arcaro MJ, Pinsk MA, Kastner S. 2015. The Anatomical and Functional Organization of the Human Visual Pulvinar. Journal of Neuroscience **35**:9848–9871. DOI: https://doi.org/10.1523/JNEUROSCI.1575-14.2015, PMID: 26156987

Badarnee M, Wen Z, Hammoud MZ, Glimcher P, Cain CK, Milad MR. 2025. Intersect between brain mechanisms of conditioned threat, active avoidance, and reward. Communications Psychology **3**:32. DOI: https://doi.org/10.1038/S44271-025-00197-7

Battaglia S, Garofalo S, di Pellegrino G, Starita F. 2020. Revaluing the Role of vmPFC in the Acquisition of Pavlovian Threat Conditioning in Humans. Journal of Neuroscience **40**:8491–8500. DOI: https://doi.org/10.1523/JNEUROSCI.0304-20.2020, PMID: 33020217

Behrens TEJ, Johansen-Berg H, Woolrich MW, Smith SM, Wheeler-Kingshott CAM, Boulby PA, Barker GJ, Sillery EL, Sheehan K, Ciccarelli O, Thompson AJ, Brady JM, Matthews PM. 2003. Non-invasive mapping of connections between human

thalamus and cortex using diffusion imaging. Nature Neuroscience **6**:750–757. DOI: https://doi.org/10.1038/NN1075;KWRD=BIOMEDICINE, PMID: 12808459
Berman RA, Wurtz RH. 2010. Functional Identification of a Pulvinar Path from Superior Colliculus to Cortical Area MT. Journal of Neuroscience **30**:6342–6354. DOI: https://doi.org/10.1523/JNEUROSCI.6176-09.2010, PMID: 20445060
Berman RA, Wurtz RH. 2008. EXPLORING THE PULVINAR PATH TO VISUAL CORTEX. Progress in brain research **171**:467. DOI: https://doi.org/10.1016/S0079-6123(08)00668-7, PMID: 18718342
Bertini C, Pietrelli M, Braghittoni D, Làdavas E. 2018. Pulvinar lesions disrupt fear-related implicit visual processing in hemianopic patients. Frontiers in Psychology **9**:412050. DOI: https://doi.org/10.3389/FPSYG.2018.02329/BIBTEX
Biesanz JC, Falk CF, Savalei V. 2010. Assessing Mediational Models: Testing and Interval Estimation for Indirect Effects. http://dx.doi.org/10.1080/00273171.2010.498292 **45**:661–701. DOI: https://doi.org/10.1080/00273171.2010.498292
Bouton ME. 2002. Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. Biological Psychiatry **52**:976–986. DOI: https://doi.org/10.1016/S0006-3223(02)01546-9, PMID: 12437938
Bridge H, Leopold DA, Bourne JA. 2015. Adaptive pulvinar circuitry supports visual cognition. Trends in cognitive sciences **20**:146. DOI: https://doi.org/10.1016/J.TICS.2015.10.003, PMID: 26553222
Carr JA. 2015. I'll take the low road: The evolutionary underpinnings of visually triggered fear. Frontiers in Neuroscience **9**:165437. DOI: https://doi.org/10.3389/FNINS.2015.00414/BIBTEX
Casanova C, Chalupa LM. 2023. The dorsal lateral geniculate nucleus and the pulvinar as essential partners for visual cortical functions. Frontiers in Neuroscience **17**:1258393. DOI: https://doi.org/10.3389/FNINS.2023.1258393/BIBTEX
Cortes N, Ladret HJ, Abbas-Farishta R, Casanova C. 2024. The pulvinar as a hub of visual processing and cortical integration. Trends in Neurosciences **47**:120–134. DOI: https://doi.org/10.1016/J.TINS.2023.11.008, PMID: 38143202
Costa M, Lozano-Soldevilla D, Gil-Nagel A, Toledano R, Oehrn CR, Kunz L, Yebra M, Mendez-Bertolo C, Stieglitz L, Sarnthein J, Axmacher N, Moratti S, Strange BA. 2022. Aversive memory formation in humans involves an amygdala-hippocampus phase code. Nature Communications 2022 13:1 **13**:1–16. DOI: https://doi.org/10.1038/s41467-022-33828-2, PMID: 36302909
Epskamp S, Borsboom D, Fried EI. 2018. Estimating psychological networks and their accuracy: A tutorial paper. Behavior research methods **50**:195–212. DOI: https://doi.org/10.3758/s13428-017-0862-1, PMID: 28342071
Esteban O, Ciric R, Finc K, Blair RW, Markiewicz CJ, Moodie CA, Kent JD, Goncalves M, DuPre E, Gomez DEP, Ye Z, Salo T, Valabregue R, Amlien IK, Liem F, Jacoby N, Stojić H, Cieslak M, Urchs S, Halchenko YO, Ghosh SS, De La Vega A, Yarkoni T, Wright J, Thompson WH, Poldrack RA, Gorgolewski KJ. 2020. Analysis of task-based functional MRI data preprocessed with fMRIPrep. Nature protocols **15**:2186–2202. DOI: https://doi.org/10.1038/S41596-020-0327-3, PMID: 32514178
Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack

RA, Gorgolewski KJ. 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nature methods **16**:111–116. DOI: https://doi.org/10.1038/S41592-018-0235-4, PMID: 30532080

Fanselow MS, LeDoux JE. 1999. Why we think plasticity underlying pavlovian fear conditioning occurs in the basolateral amygdala. Neuron **23**:229–232. DOI: https://doi.org/10.1016/S0896-6273(00)80775-8, PMID: 10399930

Foygel R, Drton M. 2010. Extended Bayesian Information Criteria for Gaussian Graphical Models. Advances in Neural Information Processing Systems **23**.

Fullana MA, Harrison BJ, Soriano-Mas C, Vervliet B, Cardoner N, Àvila-Parcet A, Radua J. 2015. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. Molecular Psychiatry 2016 21:4 **21**:500–508. DOI: https://doi.org/10.1038/mp.2015.88, PMID: 26122585

Griffiths BJ, Zaehle T, Repplinger S, Schmitt FC, Voges J, Hanslmayr S, Staudigl T. 2022. Rhythmic interactions between the mediodorsal thalamus and prefrontal cortex precede human visual perception. Nature Communications 2022 13:1 **13**:1–11. DOI: https://doi.org/10.1038/s41467-022-31407-z, PMID: 35768419

Halassa MM, Sherman SM. 2019. Thalamo-cortical circuit motifs: a general framework. Neuron **103**:762. DOI: https://doi.org/10.1016/J.NEURON.2019.06.005, PMID: 31487527

Halverson HE, Freeman JH. 2010. Ventral lateral geniculate input to the medial pons is necessary for visual eyeblink conditioning in rats. Learning & memory (Cold Spring Harbor, N.Y.) **17**:80–85. DOI: https://doi.org/10.1101/LM.1572710, PMID: 20154353

Herry C, Ferraguti F, Singewald N, Letzkus JJ, Ehrlich I, Lüthi A. 2010. Neuronal circuits of fear extinction. European Journal of Neuroscience **31**:599–612. DOI: https://doi.org/10.1111/J.1460-9568.2010.07101.X, PMID: 20384807

Homman-Ludiye J, Bourne JA. 2019. The medial pulvinar: function, origin and association with neurodevelopmental disorders. Journal of anatomy **235**:507–520. DOI: https://doi.org/10.1111/JOA.12932, PMID: 30657169

Hummos A, Wang BA, Drammis S, Halassa MM, Pleger B. 2022. Thalamic regulation of frontal interactions in human cognitive flexibility. PLOS Computational Biology **18**:e1010500. DOI: https://doi.org/10.1371/JOURNAL.PCBI.1010500, PMID: 36094955

Hwang K, Bertolero MA, Liu WB, D'Esposito M. 2017. The Human Thalamus Is an Integrative Hub for Functional Brain Networks. Journal of Neuroscience **37**:5594–5607. DOI: https://doi.org/10.1523/JNEUROSCI.0067-17.2017, PMID: 28450543

Hwang K, Bruss J, Tranel D, Boes AD. 2020. Network Localization of Executive Function Deficits in Patients with Focal Thalamic Lesions. Journal of cognitive neuroscience **32**:2303. DOI: https://doi.org/10.1162/JOCN_A_01628, PMID: 32902335

Ivan Pavlov – Nobel Lecture - NobelPrize.org. n.d. https://www.nobelprize.org/prizes/medicine/1904/pavlov/lecture/

JASP Team. 2024. JASP (Versions 0.18.3 and 0.19.3).

Koller K, Rafal RD, Platt A, Mitchell ND. 2019. Orienting toward threat: Contributions of a subcortical pathway transmitting retinal afferents to the amygdala via the superior

colliculus and pulvinar. Neuropsychologia **128**:78–86. DOI: https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2018.01.027, PMID: 29410291
Kolling N, Wittmann M, Rushworth MFS. 2014. Multiple neural mechanisms of decision making and their competition under changing risk pressure. Neuron **81**:1190–1202. DOI: https://doi.org/10.1016/J.NEURON.2014.01.033/ATTACHMENT/A9984A25-8961-453B-8E1A-778EAD3A0ADE/MMC1.PDF, PMID: 24607236
Konrad C, Neuhoff L, Adolph D, Goerigk S, Herbert JS, Jagusch-Poirier J, Weigelt S, Seehagen S, Schneider S. 2024. Associative learning via eyeblink conditioning differs by age from infancy to adulthood. Communications Psychology 2024 2:1 **2**:1–13. DOI: https://doi.org/10.1038/s44271-024-00176-4
Korn CW, Bach DR. 2019. Minimizing threat via heuristic and optimal policies recruits hippocampus and medial prefrontal cortex. Nature Human Behaviour 2019 3:7 **3**:733–745. DOI: https://doi.org/10.1038/s41562-019-0603-9, PMID: 31110338
Korn CW, Bach DR. 2018. Heuristic and optimal policy computations in the human brain during sequential decision-making. Nature Communications 2018 9:1 **9**:1–15. DOI: https://doi.org/10.1038/s41467-017-02750-3, PMID: 29362449
Kragel PA, Čeko M, Theriault J, Chen D, Satpute AB, Wald LW, Lindquist MA, Feldman Barrett L, Wager TD. 2021. A human colliculus-pulvinar-amygdala pathway encodes negative emotion. Neuron **109**:2404-2412.e5. DOI: https://doi.org/10.1016/J.NEURON.2021.06.001/ATTACHMENT/2E7063C2-7363-4DE3-9A2B-F75C55D9DD96/MMC3.PDF, PMID: 34166604
Krasne FB, Zinn R, Vissel B, Fanselow MS, Franklin Krasne CB, Howland-Rose V, Battersby D, King D. 2021. Extinction and discrimination in a Bayesian model of context fear conditioning (BaconX). Hippocampus **31**:790–814. DOI: https://doi.org/10.1002/HIPO.23298, PMID: 33452843
Le Q Van, Isbell LA, Matsumoto J, Nguyen M, Hori E, Maior RS, Tomaz C, Tran AH, Ono T, Nishijo H. 2013. Pulvinar neurons reveal neurobiological evidence of past selection for rapid detection of snakes. Proceedings of the National Academy of Sciences of the United States of America **110**:19000–19005. DOI: https://doi.org/10.1073/PNAS.1312648110/SUPPL_FILE/PNAS.201312648SI.PDF, PMID: 24167268
LeDoux JE, Pine DS. 2016. Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. https://doi.org/10.1176/appi.ajp.2016.16030353 **173**:1083–1093. DOI: https://doi.org/10.1176/APPI.AJP.2016.16030353
Lee S, Shin HS. 2016. The role of mediodorsal thalamic nucleus in fear extinction. Journal of Analytical Science and Technology **7**:1–5. DOI: https://doi.org/10.1186/S40543-016-0093-6/FIGURES/1
Lee Sukchan, Ahmed T, Lee Soojung, Kim H, Choi S, Kim DS, Kim SJ, Cho J, Shin HS. 2011. Bidirectional modulation of fear extinction by mediodorsal thalamic firing in mice. Nature Neuroscience 2011 15:2 **15**:308–314. DOI: https://doi.org/10.1038/nn.2999, PMID: 22197828
Li K, Fan L, Cui Y, Wei X, He Y, Yang J, Lu Y, Li W, Shi W, Cao L, Cheng L, Li A, You B, Jiang T. 2022. The human mediodorsal thalamus: Organization, connectivity, and function. NeuroImage **249**:118876. DOI: https://doi.org/10.1016/J.NEUROIMAGE.2022.118876, PMID: 34998970

Li XB, Inoue T, Nakagawa S, Koyama T. 2004. Effect of mediodorsal thalamic nucleus lesion on contextual fear conditioning in rats. Brain Research **1008**:261–272. DOI: https://doi.org/10.1016/J.BRAINRES.2004.02.038, PMID: 15145764

Lithari C, Moratti S, Weisz N. 2015. Thalamocortical interactions underlying visual fear conditioning in humans. Human Brain Mapping **36**:4592. DOI: https://doi.org/10.1002/HBM.22940, PMID: 26287369

Mahoney HL, Schmidt TM. 2024. The cognitive impact of light: illuminating ipRGC circuit mechanisms. Nature Reviews Neuroscience 2024 25:3 **25**:159–175. DOI: https://doi.org/10.1038/s41583-023-00788-5, PMID: 38279030

Maren S. 2001. Nuerobiology of Pavlovian fear conditioning. Annual Review of Neuroscience **24**:897–931. DOI: https://doi.org/10.1146/ANNUREV.NEURO.24.1.897/CITE/REFWORKS, PMID: 11520922

Maren S, Phan KL, Liberzon I. 2013. The contextual brain: implications for fear conditioning, extinction and psychopathology. Nature Reviews Neuroscience 2013 14:6 **14**:417–428. DOI: https://doi.org/10.1038/nrn3492, PMID: 23635870

Maren S, Quirk GJ. 2004. Neuronal signalling of fear memory. Nature Reviews Neuroscience 2004 5:11 **5**:844–852. DOI: https://doi.org/10.1038/nrn1535, PMID: 15496862

Marin MF, Hammoud MZ, Klumpp H, Simon NM, Milad MR. 2020. Multimodal Categorical and Dimensional Approaches to Understanding Threat Conditioning and Its Extinction in Individuals With Anxiety Disorders. JAMA Psychiatry **77**:618–627. DOI: https://doi.org/10.1001/JAMAPSYCHIATRY.2019.4833, PMID: 32022832

McFadyen J, Mattingley JB, Garrido MI. 2019. An afferent white matter pathway from the pulvinar to the amygdala facilitates fear recognition. eLife **8**. DOI: https://doi.org/10.7554/ELIFE.40766, PMID: 30648533

Milad MR, Furtak SC, Greenberg JL, Keshaviah A, Im JJ, Falkenstein MJ, Jenike M, Rauch SL, Wilhelm S. 2013. Deficits in Conditioned Fear Extinction in Obsessive-Compulsive Disorder and Neurobiological Changes in the Fear Circuit. JAMA Psychiatry **70**:608–618. DOI: https://doi.org/10.1001/JAMAPSYCHIATRY.2013.914

Milad MR, Pitman RK, Ellis CB, Gold AL, Shin LM, Lasko NB, Zeidan MA, Handwerger K, Orr SP, Rauch SL. 2009. Neurobiological Basis of Failure to Recall Extinction Memory in Posttraumatic Stress Disorder. Biological Psychiatry **66**:1075–1082. DOI: https://doi.org/10.1016/J.BIOPSYCH.2009.06.026, PMID: 19748076

Milad MR, Quirk GJ. 2012. Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. Annual review of psychology **63**:129. DOI: https://doi.org/10.1146/ANNUREV.PSYCH.121208.131631, PMID: 22129456

Milad MR, Quirk GJ. 2002. Neurons in medial prefrontal cortex signal memory for fear extinction. Nature 2002 420:6911 **420**:70–74. DOI: https://doi.org/10.1038/nature01138, PMID: 12422216

Milad MR, Rosenbaum BL, Simon NM. 2014. Neuroscience of fear extinction: Implications for assessment and treatment of fear-based and anxiety related disorders. Behaviour Research and Therapy **62**:17–23. DOI: https://doi.org/10.1016/J.BRAT.2014.08.006, PMID: 25204715

Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL. 2007. Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus

in concert. Biological psychiatry **62**:446–454. DOI: https://doi.org/10.1016/J.BIOPSYCH.2006.10.011, PMID: 17217927
Mukherjee A, Lam NH, Wimmer RD, Halassa MM. 2021. Thalamic circuits for independent control of prefrontal signal and noise. Nature 2021 600:7887 **600**:100–104. DOI: https://doi.org/10.1038/s41586-021-04056-3, PMID: 34614503
Nejati V, Majdi R, Salehinejad MA, Nitsche MA. 2021. The role of dorsolateral and ventromedial prefrontal cortex in the processing of emotional dimensions. Scientific Reports 2021 11:1 **11**:1–12. DOI: https://doi.org/10.1038/s41598-021-81454-7, PMID: 33479323
Nieto-Castanon A, Whitfield-Gabrieli S. 2022. CONN functional connectivity toolbox: RRID SCR_009550, release 22. CONN functional connectivity toolbox: RRID SCR_009550, release 22. DOI: https://doi.org/10.56441/HILBERTPRESS.2246.5840
O'Connor DH, Fukui MM, Pinsk MA, Kastner S. 2002. Attention modulates responses in the human lateral geniculate nucleus. Nature Neuroscience 2002 5:11 **5**:1203–1209. DOI: https://doi.org/10.1038/nn957, PMID: 12379861
Paydar A, Lee B, Gangadharan G, Lee S, Hwang EM, Shin HS. 2014. Extrasynaptic GABAA receptors in mediodorsal thalamic nucleus modulate fear extinction learning. Molecular Brain **7**:39. DOI: https://doi.org/10.1186/1756-6606-7-39, PMID: 24886120
Penzo MA, Robert V, Tucciarone J, De Bundel D, Wang M, Van Aelst L, Darvas M, Parada LF, Palmiter RD, He M, Huang ZJ, Li B. 2015. The paraventricular thalamus controls a central amygdala fear circuit. Nature **519**:455–459. DOI: https://doi.org/10.1038/NATURE13978, PMID: 25600269
Pessoa L, Adolphs R. 2010. Emotion processing and the amygdala: from a "low road" to "many roads" of evaluating biological significance. Nature Reviews Neuroscience 2010 11:11 **11**:773–782. DOI: https://doi.org/10.1038/nrn2920, PMID: 20959860
Phelps EA, LeDoux JE. 2005. Contributions of the amygdala to emotion processing: From animal models to human behavior. Neuron **48**:175–187. DOI: https://doi.org/10.1016/J.NEURON.2005.09.025/ASSET/31E14E48-ACAC-4132-ACD2-31705A61F19C/MAIN.ASSETS/GR4.JPG, PMID: 16242399
Rafal RD, Koller K, Bultitude JH, Mullins P, Ward R, Mitchell AS, Bell AH. 2015. Connectivity between the superior colliculus and the amygdala in humans and macaque monkeys: virtual dissection with probabilistic DTI tractography. Journal of neurophysiology **114**:1947–1962. DOI: https://doi.org/10.1152/JN.01016.2014, PMID: 26224780
Ramanathan KR, Jin J, Giustino TF, Payne MR, Maren S. 2018. Prefrontal projections to the thalamic nucleus reuniens mediate fear extinction. Nature Communications 2018 9:1 **9**:1–12. DOI: https://doi.org/10.1038/s41467-018-06970-z, PMID: 30375397
Ramanathan KR, Maren S. 2019. Nucleus reuniens mediates the extinction of contextual fear conditioning. Behavioural Brain Research **374**:112114. DOI: https://doi.org/10.1016/J.BBR.2019.112114, PMID: 31351844
Ratigan HC, Krishnan S, Smith S, Sheffield MEJ. 2023. A thalamic-hippocampal CA1 signal for contextual fear memory suppression, extinction, and discrimination.

Nature Communications 2023 14:1 **14**:1–17. DOI: https://doi.org/10.1038/s41467-023-42429-6, PMID: 37875465
Rolls ET, Huang CC, Lin CP, Feng J, Joliot M. 2020. Automated anatomical labelling atlas 3. NeuroImage **206**. DOI: https://doi.org/10.1016/J.NEUROIMAGE.2019.116189, PMID: 31521825
Rosseel Y. 2012. lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software **48**:1–36. DOI: https://doi.org/10.18637/JSS.V048.I02
Saalmann YB, Kastner S. 2011. Cognitive and Perceptual Functions of the Visual Thalamus. Neuron **71**:209. DOI: https://doi.org/10.1016/J.NEURON.2011.06.027, PMID: 21791281
Saalmann YB, Kastner S. 2009. Gain control in the visual thalamus during perception and cognition. Current opinion in neurobiology **19**:408. DOI: https://doi.org/10.1016/J.CONB.2009.05.007, PMID: 19556121
Salay LD, Huberman AD. 2021. Divergent outputs of the ventral lateral geniculate nucleus mediate visually evoked defensive behaviors. Cell Reports **37**. DOI: https://doi.org/10.1016/j.celrep.2021.109792, PMID: 34610302
Sherman SM. 2016. Thalamus plays a central role in ongoing cortical functioning. Nature Neuroscience 2016 19:4 **19**:533–541. DOI: https://doi.org/10.1038/nn.4269, PMID: 27021938
Sherman SM. 2007. The thalamus is more than just a relay. Current opinion in neurobiology **17**:417–422. DOI: https://doi.org/10.1016/J.CONB.2007.07.003, PMID: 17707635
Sherman SMurray, Guillery RW. 2006. Exploring the thalamus and its role in cortical function. 484.
Shi C, Davis M. 2001. Visual pathways involved in fear conditioning measured with fear-potentiated startle: behavioral and anatomic studies. The Journal of neuroscience□: the official journal of the Society for Neuroscience **21**:9844–9855. DOI: https://doi.org/10.1523/JNEUROSCI.21-24-09844.2001, PMID: 11739592
Silverstein DN, Ingvar M. 2015. A multi-pathway hypothesis for human visual fear signaling. Frontiers in Systems Neuroscience **9**:137726. DOI: https://doi.org/10.3389/FNSYS.2015.00101/PDF
Steinmetz AB, Buss EW, Freeman JH. 2013. Inactivation of the ventral lateral geniculate and nucleus of the optic tract impairs retention of visual eyeblink conditioning. Behavioral Neuroscience **127**:690–693. DOI: https://doi.org/10.1037/A0033729, PMID: 23978151
Takakuwa N, Isa K, Onoe H, Takahashi J, Isa T. 2021. Contribution of the Pulvinar and Lateral Geniculate Nucleus to the Control of Visually Guided Saccades in Blindsight Monkeys. Journal of Neuroscience **41**:1755–1768. DOI: https://doi.org/10.1523/JNEUROSCI.2293-20.2020, PMID: 33443074
Totty MS, Tuna T, Ramanathan KR, Jin J, Peters SE, Maren S. 2023. Thalamic nucleus reuniens coordinates prefrontal-hippocampal synchrony to suppress extinguished fear. Nature Communications 2023 14:1 **14**:1–12. DOI: https://doi.org/10.1038/s41467-023-42315-1, PMID: 37848425
Tovote P, Fadok JP, Lüthi A. 2015. Neuronal circuits for fear and anxiety. Nature Reviews Neuroscience 2015 16:6 **16**:317–331. DOI: https://doi.org/10.1038/nrn3945, PMID: 25991441

Tuna T, Totty MS, Badarnee M, Mourão FAG, Peters S, Milad MR, Maren S. 2025. Associative coding of conditioned fear in the thalamic nucleus reuniens in rodents and humans. Communications Biology 2025 8:1 **8**:1–14. DOI: https://doi.org/10.1038/s42003-025-08580-0, PMID: 40750836

Vuilleumier P, Armony JL, Driver J, Dolan RJ. 2003. Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nature Neuroscience 2003 6:6 **6**:624–631. DOI: https://doi.org/10.1038/nn1057, PMID: 12740580

Ward R, Danziger S, Bamford S. 2005. Response to visual threat following damage to the pulvinar. Current biology□: CB **15**:571–573. DOI: https://doi.org/10.1016/J.CUB.2005.01.056, PMID: 15797028

Wei P, Liu N, Zhang Z, Liu X, Tang Y, He X, Wu B, Zhou Z, Liu Y, Li J, Zhang Y, Zhou X, Xu L, Chen L, Bi G, Hu X, Xu F, Wang L. 2015. Processing of visually evoked innate fear by a non-canonical thalamic pathway. Nature Communications 2015 6:1 **6**:1–13. DOI: https://doi.org/10.1038/ncomms7756, PMID: 25854147

Wen Z, Pace-Schott EF, Lazar SW, Rosén J, Åhs F, Phelps EA, LeDoux JE, Milad MR. 2024. Distributed neural representations of conditioned threat in the human brain. Nature Communications 2024 15:1 **15**:1–14. DOI: https://doi.org/10.1038/s41467-024-46508-0, PMID: 38472184

Wen Z, Raio CM, Pace-Schott EF, Lazar SW, LeDoux JE, Phelps EA, Milad MR. 2022. Temporally and anatomically specific contributions of the human amygdala to threat and safety learning. Proceedings of the National Academy of Sciences of the United States of America **119**:e2204066119. DOI: https://doi.org/10.1073/PNAS.2204066119/SUPPL_FILE/PNAS.2204066119.SAPP.PDF, PMID: 35727981

Whitfield-Gabrieli S, Nieto-Castanon A. 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain connectivity **2**:125–141. DOI: https://doi.org/10.1089/BRAIN.2012.0073, PMID: 22642651

Whyte CJ, Redinbaugh MJ, Shine JM, Saalmann YB. 2024. Thalamic contributions to the state and contents of consciousness. Neuron **112**:1611–1625. DOI: https://doi.org/10.1016/J.NEURON.2024.04.019/ASSET/AA96B159-0E29-4D7D-B3C8-E302EA442A73/MAIN.ASSETS/GR2.JPG, PMID: 38754373

Wolff M, Halassa MM. 2024. The mediodorsal thalamus in executive control. Neuron **112**:893–908. DOI: https://doi.org/10.1016/J.NEURON.2024.01.002/ASSET/8EBBDE5D-737B-479F-966E-07AE8F092DB6/MAIN.ASSETS/GR3.JPG, PMID: 38295791

Wurtz RH, McAlonan K, Cavanaugh J, Berman RA. 2011. Thalamic pathways for active vision. Trends in Cognitive Sciences **15**:177–184. DOI: https://doi.org/10.1016/J.TICS.2011.02.004, PMID: 21414835

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. Nature Methods 2011 8:8 **8**:665–670. DOI: https://doi.org/10.1038/nmeth.1635, PMID: 21706013

Zheng C, Huang Y, Bo B, Wei L, Liang Z, Wang Z. 2020. Projection from the Anterior Cingulate Cortex to the Lateral Part of Mediodorsal Thalamus Modulates Vicarious Freezing Behavior. Neuroscience Bulletin **36**:217–229. DOI: https://doi.org/10.1007/S12264-019-00427-Z/METRICS, PMID: 31531804