Individual differences in emotional learning–driven prioritization of neutral episodic memories across temporal boundaries

Blazej M. Baczkowski[1], Michiko Sakaki[1,2], Felix Kalbe[3], Kou Murayama[1,2], and Lars Schwabe[3]

[1]Hector Research Institute of Education Sciences and Psychology, University of Tübingen

[2]Research Institute, Kochi University of Technology

[3]Department of Cognitive Psychology, University of Hamburg

Author Note

Blazej M. Baczkowski https://orcid.org/0000-0001-9827-797X

Michiko Sakaki https://orcid.org/0000-0003-1993-5765

Felix Kalbe https://orcid.org/0000-0003-4429-4373

Kou Murayama https://orcid.org/0000-0003-2902-9600

Lars Schwabe https://orcid.org/0000-0003-4429-4373

This manuscript is a preprint and has not been peer reviewed.

Correspondence concerning this article should be addressed to Blazej M. Baczkowski, Email: blazej.baczkowski@uni-tuebingen.de

Abstract

Episodic memory is shaped by event segmentation, which divides continuous experience into discrete episodes often marked by temporal gaps. Emotional learning can influence memory across these boundaries by prioritizing not only threatening events but also related information that is encoded shortly before or after the emotional episode, which in turn supports adaptive generalization. However, evidence for such cross-boundary modulation is mixed, possibly because conventional group-level analyses can mask systematic individual variability in the effects of emotional learning on nearby neutral memories.

We reanalyzed data from a large-sample study (N = 285) combining emotional and episodic learning that had previously shown no consistent retroactive or proactive enhancement across event boundaries at the group level. Young adults incidentally encoded neutral images from two categories (animals, tools) across three temporally separated phases: pre-conditioning, conditioning, and post-conditioning. One category was partially reinforced with mild electric shocks during conditioning to create a category-level threat association. Memory was tested 24 hours later using a surprise old–new recognition task.

A hierarchical Bayesian recognition model with a latent-mixture component revealed two profiles of memory modulation. One showed declining memory across phases, with modest enhancement limited to items from the shock-predictive category encoded during conditioning. The other showed strongest memory for items encoded during conditioning, with substantial category-selective enhancement extending into post-conditioning. These findings suggest that proactive, but not retroactive, enhancement is most evident in individuals whose memories are strongly anchored to the emotional episode, indicating integration of threat-related information across temporal boundaries.

Keywords: episodic memory, emotional learning, threat conditioning, event segmentation, memory recognition

Individual differences in emotional learning–driven prioritization of neutral episodic memories across temporal boundaries

Introduction

Human memory is not merely a passive record of experience, but a highly adaptive system that selectively filters information to guide future behavior (Nairne et al., 2007; Shohamy & Adcock, 2010). Emotionally salient experiences, which are relevant to survival, are often prioritized for long-term storage (Talmi, 2013; Williams et al., 2022). For example, while routine aspects of a forest hike – trees, trail markers, birdsong – may fade quickly, a threatening encounter with a bear is likely to persist in memory. It may also be adaptive to remember related cues encountered nearby in time, such as the rustling of leaves heard minutes before the bear appeared or claw marks noticed on the way back home. Such cues can later serve as warning signals, even though they were encoded in safe contexts. Emotionally significant experiences may therefore enhance episodic memory for both the emotional and neutral events, supporting adaptive generalization – the capacity to infer danger from memories acquired without direct aversive experience (Baczkowski et al., 2023).

Emotional learning is known to strengthen episodic memory for the aversive episode itself. The bear example illustrates that a previously neutral cue can become threatening, as described by Pavlovian threat conditioning (LeDoux, 2000; Rescorla, 1988). This associative learning not only links cues to aversive outcomes but also enhances episodic memory for the aversive experience via the mechanisms of prediction error and arousal-related modulation of amygdala–hippocampus interactions (de Voogd et al., 2016b; Kalbe & Schwabe, 2021b; Rouhani et al., 2023). Yet such enhancement may not extend to neutral events occurring close in time but outside the emotional context. This limitation may arise from event segmentation – the cognitive parsing of continuous experience into discrete episodes (Kurby & Zacks, 2008; Zacks & Swallow, 2007), often marked by perceptual shifts such as temporal gaps (Clewett et al., 2025; Clewett & Davachi, 2017; Ezzyat & Davachi, 2010). By isolating threat from competing experiences of safety (Dunsmoor et al., 2018; Ezzyat & Davachi, 2021), event segmentation may guard against overgeneralization but also constrain the prioritization effects of emotional learning on neutral events occurring in close temporal proximity.

Despite the constraints imposed by event segmentation, several complementary mechanisms have been proposed to extend emotional memory benefits across temporal boundaries (Clewett et al., 2019). The first is behavioral tagging, whereby a weak memory is stabilized if it occurs near a salient event. This relies on synaptic-level tag-and-capture processes (Ballarini et al., 2009; Frey & Morris, 1997), in which strong emotional experiences trigger neuromodulatory activity and protein synthesis that consolidate both the emotional and preceding neutral content (Dunsmoor et al., 2022). A second route is post-encoding reactivation (Hermans et al., 2016; Loock et al., 2025), in which emotional events partially reinstate earlier related neutral representations, strengthening their consolidation through persistent hippocampal–cortical activity or replay (Staresina et al., 2013; Tambini & Davachi, 2013). A third route, in contrast, enhances memory for subsequent events through sustained arousal-related brain states, characterized by prolonged amygdala–hippocampus activity (Clewett & McClay, 2024; Tambini et al., 2016). Together, these processes may enable emotional learning to prioritize memory for temporally proximal but separated experiences, thereby supporting adaptive generalization.

However, evidence for such cross-boundary enhancement is mixed. Dunsmoor et al. (2015) reported that emotional learning prioritized memory for threat-related information encoded in a neutral context shortly before or after conditioning. In their design combining emotional and episodic learning, participants incidentally encoded non-repeating, neutral images of basic-level exemplars from two semantic categories – animals and tools – across three phases: pre-conditioning, conditioning, and post-conditioning (Figure 1, panel a). Short breaks between phases marked event boundaries, operationalizing each phase as a separate memory episode. During conditioning, one category was partially reinforced with mild electric shocks, establishing a Pavlovian association at the category rather than the item level (Dunsmoor et al., 2012, 2014). Category-level association enhanced memory for items from the shock-predictive category encoded during conditioning irrespective of direct reinforcement, leading to within-category mnemonic generalization. Critically, neutral items from the shock-predictive category that had been encoded without risk of shock before or after conditioning were remembered better 24 hours later, as shown by a surprise old–new recognition test.

In contrast, Kalbe and Schwabe (2021a) closely replicated the emotional–episodic learning paradigm in four studies with 285 participants in total but found no consistent evidence of retroactive or proactive cross-boundary memory strengthening, despite robust within-phase emotional memory enhancement. One explanation for this discrepancy is that averaging across participants may mask meaningful individual differences in how emotional learning affects memory – effects that may be present in some individuals but not others (see also Bolger et al., 2019; Olsson-Collentine et al., 2020). Individual variability could stem from differences in learning rate, physiological arousal, or consolidation processes (de Voogd et al., 2016a; Kalbe & Schwabe, 2020), which may in turn determine whether mechanisms such as behavioral tagging, post-encoding reactivation, or sustained arousal-related brain states enable emotional memory benefits to extend across temporal boundaries.

To examine this possibility, we reanalyzed the Kalbe and Schwabe (2021a) dataset using a model-based framework that captures individual-level estimates of recognition memory and reveals hidden patterns that group averages can obscure. We adopted the two-high-threshold (2HT) recognition model (Snodgrass & Corwin, 1988), which distinguishes memory-based recognition from guessing, to ensure theoretical continuity with prior work (Dunsmoor et al., 2015). Rather than relying on corrected recognition scores – a summary metric that assumes a correspondence between hit and false alarm rates – we implemented the 2HT model in its original probabilistic decision-tree form (Figure 1, panel b). This allowed us to accommodate the asymmetry in our design between targets (varied by both phase and category) and lures (varied only by category).

We first fit the 2HT model in a hierarchical Bayesian framework to estimate recognition memory at both the group and individual levels. Next, we applied data-driven clustering to explore patterns in individual-level recognition estimates. Finally, to test whether these patterns reflected meaningful differences in retroactive and proactive effects of Pavlovian threat conditioning on episodic memory, we extended the model with a latent-mixture component that probabilistically assigned participants to distinct memory profiles (Lee & Stark, 2023). Collectively, these analyses provided an approach more sensitive to individual variability, moving beyond binary replication outcomes and toward a more nuanced understanding of how emotional learning affects episodic memory.

Methods

Data source

The present study re-analyzed data from Kalbe and Schwabe (2021a), who conducted four studies aimed to test retroactive and proactive effects of Pavlovian threat conditioning on episodic memory. These studies closely replicated the experimental procedure of Dunsmoor et al. (2015), with minor variations in timing, instructions, and stimuli. Behavioral data were obtained from a public repository (https://osf.io/qpm3t), with additional demographic and skin conductance data provided by the corresponding author upon request.

Data collection procedures

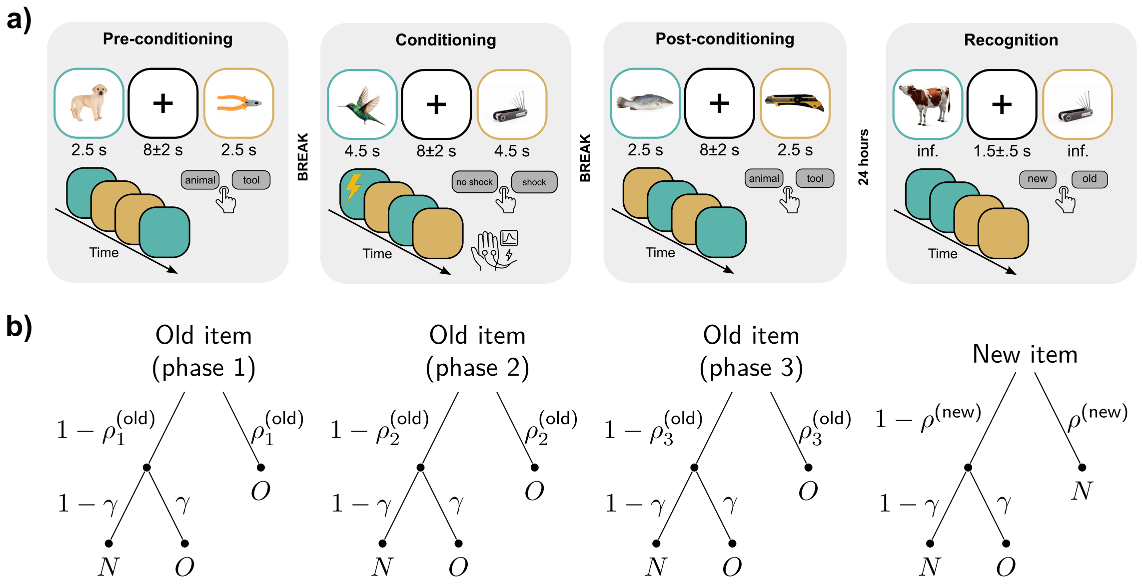
The original protocol was approved by the local ethics committee. All participants gave written informed consent and received monetary compensation. This re-analysis required no additional ethical approval. Full methods are reported in Kalbe and Schwabe (2021a).

Briefly, each study consisted of two sessions approximately 24 hours apart. In session 1, participants incidentally encoded 180 unique neutral images from two semantic categories (animals and tools) across three separate phases: pre-conditioning, conditioning, and post-conditioning. In pre- and post-conditioning, participants categorized 60 images per phase (30 per category), each presented for 2.5 seconds with interstimulus intervals of 6 ± 2 seconds.

During conditioning, participants viewed another set of 60 images (30 per category) each shown for 4.5 seconds, followed by an 8 ± 2 second fixation, while electrodes for measuring skin conductance and delivering unpleasant electric shocks were attached. Images from one category (CS+) were predictive of the shock (US) at a 2/3 contingency, while images from the other category (CS-) were never paired. The assignment of semantic categories to CS+ and CS- conditions was pseudo-randomized across participants. On each

Figure 1

Experimental design and decision tree of old-new recognition judgments according to two-high-threshold model (2HT).



Note. Panel a: In each phase, participants incidentally encoded 30 unique images each of animals and tools. During pre- and post-conditioning phases, they categorized images as animals or tools; in the conditioning phase, one category probabilistically predicted mild electric shocks, and participants indicated their shock expectancy. About 24 hours later, they completed a surprise old-new recognition test including all previously seen images and an equal number of new ones (confidence ratings not shown). Colored borders are illustrative. Visual elements adapted from `flaticon.com` (designed by flaticon, juicy_fish, Smashicons) and `freepik.com` (designed by freepik, evening_tao, brgfx). Panel b: The two-high-threshold (2HT) model formalizes old–new recognition judgments (deciding old, O vs. new, N) as a probabilistic decision tree involving two cognitive states: detection – a memory-based recognition (ρ) – and uncertainty leading to guessing old (γ). Items are detected as old when they exceed the old-item threshold, indicating a match of a target to a memory from a specific encoding phase (e.g., $\rho_1^{(\text{old})}$). New items are detected as new when an item does not match to any of the encoding phases ($\rho^{(\text{new})}$). Otherwise, the model assumes a guess.

trial, participants made a binary response indicating whether they expected a shock, knowing their response would not influence outcomes.

In session 2, participants completed a surprise old-new recognition test with 360 images, in which old images were mixed with an equal number of new images from each semantic category.

Participants

Data from all available participants ($N = 285$) were analysed. Demographic data were available for 277 participants (191 female), aged 18–35 years (mode = 26). Skin conductance data were available for 272 participants (185 female), all within the same age range.

Measures

Old-new recognition judgments

Binary recognition judgments (*old* vs. *new*) served as the primary outcome variable. In Study 1, judgments were collected in two steps: an old/new decision followed by a 4-point confidence rating (*very unsure*, *rather unsure*, *rather sure*, and *very sure*), with a 5-second response window. Consequently, trial counts varied slightly across participants, with 12 out of 44 participants missing only a few trials (mode = 1). In the other three studies, participants selected one of four combined memory-confidence responses (*definitely/maybe old* or *definitely/maybe new*) without a time limit. For analysis, responses were binarized (*old* vs. *new*) by collapsing confidence levels, following prior work (Dunsmoor et al., 2015; Kalbe & Schwabe, 2021a).

Data analysis

Before model fitting, we performed data quality and reproducibility checks. We visually inspected hits and false alarms (FA) for each participant and reproduced key analyses from Kalbe and Schwabe (2021a), including overall performance, repeated-measures ANOVA, and paired t-tests on corrected recognition scores (hit rate minus FA rate), which are reported in the Supplement.

After confirming data integrity, we fit a Bayesian 2HT model to old-new recognition judgments. The baseline model assumed a single generative process for all participants and formed the basis for participant clustering and a latent mixture extension for subgroup testing. Graphical representations of both models, including the data-generating processes and all prior distributions, are provided in Supplementary Figure S1.

Bayesian two-high-threshold (2HT) model of old-new recognition judgments

The 2HT model predicts the number of hits ($y_{ic}^{(\text{old})}$) and FAs ($y_{id}^{(\text{new})}$) using binomial likelihoods, where i indexes participants. Hits were collected across six encoding conditions: $c \in \{1, 2, 3\}$ for CS- and $c \in \{4, 5, 6\}$ for CS+ items seen pre-, during, and post-conditioning, respectively. FAs were collected by semantic category only: $d = 1$ for CS– and $d = 2$ for CS+.

$$y_{ic}^{(\text{old})} \sim \text{Binomial}(n_{ic}^{(\text{old})}, \theta_{ic}^{(\text{old})})$$
$$y_{id}^{(\text{new})} \sim \text{Binomial}(n_{id}^{(\text{new})}, \theta_{id}^{(\text{new})}),$$

where $\theta_{ic}^{(\text{old})}$ and $\theta_{id}^{(\text{new})}$ denote the probabilities of hits and FAs, respectively.

The probability of hits arises from either detection of an old item from condition c ($\rho_{ic}^{(\text{old})}$) or, if detection fails, a guess based on the item's semantic category (γ_{id}):

$$\theta_{ic}^{(\text{old})} = \begin{cases} \rho_{ic}^{(\text{old})} + (1 - \rho_{ic}^{(\text{old})}) \cdot \gamma_{i1}, & \text{if } c \in \{1, 2, 3\} \\ \rho_{ic}^{(\text{old})} + (1 - \rho_{ic}^{(\text{old})}) \cdot \gamma_{i2}, & \text{if } c \in \{4, 5, 6\}. \end{cases}$$

The probability of FAs arises from guessing *old* after detection has failed, i.e., when participant fails to recognize that an item was never seen during the experiment:

$$\theta_{id}^{(\text{new})} = (1 - \rho_{id}^{(\text{new})}) \cdot \gamma_{id}, \text{ for } d \in \{1, 2\}.$$

To ensure identifiability, we adopted the standard 2HT assumption that the probability of detecting old items (i.e., identifying them based on a memory match) equals the probability of detecting new items (i.e., identifying them based on the absence of a memory match):

$$\rho^{(\text{new})} = \rho^{(\text{old})}.$$

In our design, old items were tied to specific encoding phases, yielding separate hit rates for each phase and allowing estimation of phase-specific detection probabilities (i.e., $\rho_1^{(\text{old})}, \rho_2^{(\text{old})}, \rho_3^{(\text{old})}$). New items, in contrast, were tied only to a semantic category and had to be identified as *new* by ruling out matches across all encoding phases. To connect new-item detection to old-item detection, we assumed – consistent with the design logic – that underlying memories are tied to specific encoding phases and operate independently across phases. This assumption ensures model identifiability and allows us to express new-item detection as the product of phase-specific detection probabilities. This way, overall new-item detection arises from the joint contribution of memory for each phase (i.e., correctly identifying a new item as *new* requires detecting the absence of a memory match across all three phases):

$$\rho_{id}^{(\text{new})} = \begin{cases} \rho_{i1}^{(\text{old})}\rho_{i2}^{(\text{old})}\rho_{i3}^{(\text{old})} & \text{if } d = 1, \\ \rho_{i4}^{(\text{old})}\rho_{i5}^{(\text{old})}\rho_{i6}^{(\text{old})} & \text{if } d = 2. \end{cases}$$

A detection failure occurs when the absence of a memory match is not detected in one or more phases. For example, a participant may correctly identify that a new item does not match memories from the pre- and post-conditioning phases but fail to do so for the conditioning phase, resulting in a detection failure. The probability of a detection failure is given by the complement:

$$1 - \rho_{id}^{(\text{new})} = \begin{cases} 1 - \rho_{i1}^{(\text{old})}\rho_{i2}^{(\text{old})}\rho_{i3}^{(\text{old})}, & \text{if } d = 1 \\ 1 - \rho_{i4}^{(\text{old})}\rho_{i5}^{(\text{old})}\rho_{i6}^{(\text{old})}, & \text{if } d = 2. \end{cases}$$

Accordingly, weak memory for even a single phase may result in a detection failure, increasing reliance on guessing.

Assuming independence, participant-level memory recognition ρ_{ic} and guessing γ_{id} probabilities are expressed as outcomes of a linear combination of the effects of interest, encoded in a design matrix and its associated weights, with additional adjustments for procedural variations across studies and for differences related to semantic category.

Specifically, memory recognition probabilities are defined:

$$\text{logit}(\rho_{ic}^{(\text{old})}) = \boldsymbol{X}_c^{(\beta)} \cdot (\boldsymbol{\beta}_i + \boldsymbol{u}_{[study]}^{(\beta)}) + w_c \cdot \eta^{(\rho)},$$

where $\boldsymbol{X}_c^{(\beta)} \in \mathbb{R}^{1\times 6}$ is a row of the design matrix corresponding to condition c, $\boldsymbol{\beta}_i \in \mathbb{R}^{6\times 1}$ are weights, $\boldsymbol{u}_{[study]}^{(\beta)} \in \mathbb{R}^{6\times 1}$ capture adjustments for between-study variability, and $\eta^{(\rho)} \in \mathbb{R}$

accounts for the semantic category effects (tools vs. animals). The design matrix $\boldsymbol{X}^{(\beta)}$ and its associated hypothesis matrix $\boldsymbol{H}^{(\beta)} = inv(\boldsymbol{X}^{(\beta)})$ isolates the grand mean (β_1), phase comparisons (β_2: phase 2 vs. phase 1; β_3: phase 3 vs. phase 2), and CS+ vs. CS– contrasts within each phase (β_4–β_6 for pre-, during-, and post-conditioning, respectively):

$$\boldsymbol{H}^{(\beta)} = \begin{array}{c} \\ phase_1^{CS-} \\ phase_2^{CS-} \\ phase_3^{CS-} \\ phase_1^{CS+} \\ phase_2^{CS+} \\ phase_3^{CS+} \end{array} \begin{array}{c} \begin{array}{cccccc} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 \end{array} \\ \begin{bmatrix} 1/6 & -1 & 0 & -1 & 0 & 0 \\ 1/6 & 1 & -1 & 0 & -1 & 0 \\ 1/6 & 0 & 1 & 0 & 0 & -1 \\ 1/6 & -1 & 0 & 1 & 0 & 0 \\ 1/6 & 1 & -1 & 0 & 1 & 0 \\ 1/6 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Given the *logit* space of the model, individual weights $\boldsymbol{\beta}_i$ represent differences on a log-odds scale (excluding the grand mean). When these log-odds are exponentiated, they yield odds ratios – an effect size that quantifies how much more likely an event is in one condition compared to another. For example, a log-odds of 0.18 implies that the odds of correct recognition in the CS+ condition are $e^{0.18} \approx 1.2$ times higher than in the CS- condition. Individual weights $\boldsymbol{\beta}_i$ are drawn from a multivariate normal with population-level means $\boldsymbol{\mu}^{(\beta)}$ and covariance $\boldsymbol{\Sigma}^{(\beta)}$. We use weakly informative priors: Normal(0, 1) for most parameters; Normal(–0.5, 1.5) for the grand mean tailored based on prior predictive checks of the overall hits; Lewandowski-Kurowicka-Joe (shape = 2) and Gamma (shape = 5, rate = 10) distributions for the correlation matrix and standard deviations.

Guessing was modeled similarly:

$$\text{logit}(\gamma_{id}) = \boldsymbol{X}_d^{(\alpha)} \cdot (\boldsymbol{\alpha}_i + \boldsymbol{u}_{[study]}^{(\alpha)}) + w_d \cdot \eta^{\gamma},$$

with hypothesis matrix:

$$\boldsymbol{H}^{(\alpha)} = \begin{array}{c} \\ CS- \\ CS+ \end{array} \begin{array}{c} \begin{array}{cc} \alpha_1 & \alpha_2 \end{array} \\ \begin{bmatrix} 1/2 & -1 \\ 1/2 & 1 \end{bmatrix} \end{array}.$$

Here, α_1 represents overall guessing and α_2 the CS+ vs. CS– difference. Individual weights $\boldsymbol{\alpha}_i$ are again drawn from a multivariate normal, with priors matching those in the recognition component: Normal(0, 1) for most parameters and Normal(–0.8, 1.5) for the grand mean tailored according to prior predictive checks of the overall FAs.

Although we chose the 2HT model for theoretical continuity with prior work (Dunsmoor et al., 2015), we also assessed its robustness by comparing it with two alternative accounts of old–new responses: the one-high-threshold (1HT) model and a signal detection theory (SDT) model (see the Supplement). We compared models using 5-fold cross-validation stratified by subject and study (i.e., the individual experiments comprising the dataset). Each model was trained on 80% of the data and tested on the remaining 20%, with out-of-sample predictive performance measured via expected log predictive density (ELPD) via the `loo` R package (Vehtari et al., 2024). Comparisons were based on differences in ELPD, with higher values indicating better fit; following common guidelines (Vehtari et al., 2016), differences exceeding twice their standard error were considered meaningful.

Clustering individual variability in memory recognition

To identify clusters of participants based on posterior memory recognition estimates (i.e., $\boldsymbol{\beta}$ weights), we applied Kohonen self-organizing maps (SOMs) via the `kohonen` R package (Kohonen, 1982; Wehrens & Kruisselbrink, 2018). SOMs project high-dimensional data onto a 2D grid while preserving topological relationships, such that similar input vectors are mapped to nearby nodes. We trained a 7x8 hexagonal SOM (no edge wrapping) for 5000 iterations on scaled posterior means, representing each participant as a 6D $\boldsymbol{\beta}$ vector. Each vector was mapped to its best matching unit, and all SOM nodes were populated, indicating a good fit. We then computed a Euclidean distance matrix between node vectors and applied k-means clustering to group topologically adjacent nodes. The silhouette method indicated two optimal clusters. This solution was validated by directly clustering the original $\boldsymbol{\beta}$ vectors, which again supported a two-cluster structure.

Bayesian two-high-threshold (2HT) model with latent mixture extension

To formally assess participant heterogeneity, we extended the baseline 2HT model into a hierarchical latent mixture framework (Bartlema et al., 2014; Lee & Stark, 2023), where each participant is probabilistically assigned to one of two latent groups based on their observed data. We placed a uniform Beta prior on the group proportion (ϕ), assuming no prior preference. Each participant's recognition weight vector $\boldsymbol{\beta}_i$ is drawn from one of two multivariate normal distributions with distinct group-specific means $\boldsymbol{\mu}^{(\beta,\, z_i)}$ but a shared diagonal covariance matrix $\Sigma^{(\beta)}$:

$$z_i \sim \text{Bernoulli}(\phi),$$
$$\boldsymbol{\beta}_i \mid z_i \sim \text{MvNormal}(\boldsymbol{\mu}^{(\beta,\, z_i)}, \Sigma^{(\beta)})$$

where z_i denotes the latent group membership for participant i. This structure retains the baseline model's hierarchy while allowing for group-level differences. To interpret the latent subgroups, we tested for associations between group membership and individual differences in key indicators of Pavlovian learning, including the Rescorla-Wagner learning rate and anticipatory arousal measured via skin conductance responses (see the Supplement).

Bayesian parameter estimation and inference

Bayesian analyses were conducted in Stan (Stan Development Team, 2023) via the `rstan` R package (Stan Development Team, 2024), using the No-U-Turn Sampler (NUTS), a variant of Hamiltonian Monte Carlo (Betancourt, 2018; Hoffman & Gelman, 2014). Four MCMC chains were run for 20,000 iterations each (2,000 warm-up, thinning = 2), yielding 40,000 post-warm-up samples. Convergence was confirmed via traceplots, $\hat{R}$ statistics (Brooks & Gelman, 1998), and effective sample sizes (>10,000 for all key parameters).

Model adequacy was assessed through posterior predictive checks comparing simulated and observed grand means of hits and FAs, both overall and by study. For the latent mixture model, checks were conditioned on group membership, and generalizability of the two-group structure across studies was examined (see the Supplement).

We report posterior means and 89% highest posterior density intervals (HPDIs) in the format: mean [lower bound, upper bound]. Evidence for or against effects was quantified using the Region of Practical Equivalence (ROPE) (Etz et al., 2024; Kruschke, 2018; Makowski et al., 2019), with standard ranges applied to the log-odds (odds ratio) scale: [−0.18, 0.18] (≈ [0.85, 1.20] in odds ratio units) and to correlation coefficients: [-0.1, 0.1] (Cohen, 2013; Kruschke, 2014). We specified a ROPE of [0, 0.2] for the mixing proportion, because proportions close to zero imply that no meaningful subgroup is present.

Results

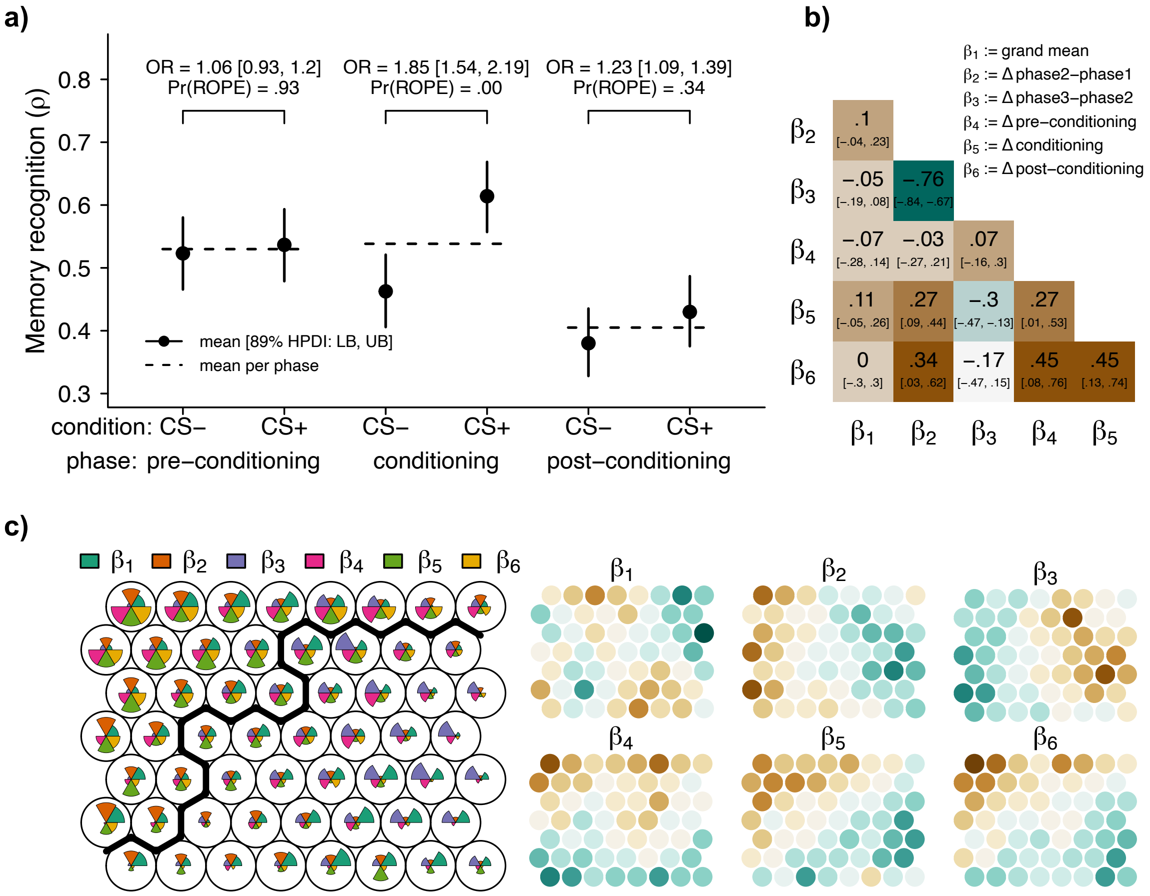
Selective memory prioritization by emotional learning varies across time and individuals

The Bayesian 2HT model yielded results largely consistent with the original publication [Figure 2; panel a; Kalbe and Schwabe (2021a)]. Recognition probability was higher for CS+ than CS- items encoded during conditioning ($\rho_{\text{CS}^+} = .61$ [.56, .67]; $\rho_{\text{CS}^-} = .46$ [.41, .52]), revealing robust category-selective memory prioritization (OR = 1.85 [1.54, 2.19], $\text{Pr}_{\text{ROPE}} = .00$), which did not generalize clearly to other phases. Recognition rates for pre-conditioning CS+ and CS− items were nearly identical ($\rho_{\text{CS}^+} = .54$ [.48, .59]; $\rho_{\text{CS}^-} = .52$ [.47, .58]; OR = 1.06 [0.93, 1.2], $\text{Pr}_{\text{ROPE}} = .93$). Post-conditioning CS+ items showed slight advantage ($\rho_{\text{CS}^+} = .43$ [.38, .49]; $\rho_{\text{CS}^-} = .38$ [.33, .44]), but overlapped with the ROPE (OR = 1.23 [1.09, 1.39], $\text{Pr}_{\text{ROPE}} = .34$), providing tentative evidence for proactive effects. These results were corroborated by both the 1HT and SDT models (see the Supplement). An ELPD-based comparison of predictive performance ranked model fit as follows: 1HT (-9252.43), 2HT (-9260.19), and then SDT (-9296.55), with differences not large enough to distinguish between models (2HT vs. 1HT: ΔELPD = 7.76, SE = 12.08, $z = 0.64$; 2HT vs. SDT: ΔELPD = 36.36, SE = 24.4, $z = 1.49$), although the evidence marginally favored the 1HT model over SDT (ΔELPD = 44.12, SE = 23.15, $z = 1.91$).

Participant-level correlations revealed individual differences in category-selective memory prioritization (Figure 2, panel b). This was most evident between conditioning and post-conditioning phases (r = .45 [.13, .74], $\text{Pr}_{\text{ROPE}} = .04$), with a tentative effect of smaller magnitude between conditioning and pre-conditioning (r = .27 [.01, .53], $\text{Pr}_{\text{ROPE}} = .14$). In other words, individuals who selectively prioritized memory for items encoded during conditioning were more likely to do so post-conditioning, and to a lesser extent, pre-conditioning. We also observed a strong negative correlation between phases regardless of CS type (r = -0.76 [-0.84, -0.67], $\text{Pr}_{\text{ROPE}} = .00$), highlighting an additional source of individual variability and motivating further exploration through clustering with Kohonen self-organizing maps. The analysis revealed a two-cluster structure (Figure 2, panel c), which we examined using individual feature maps, each reflecting similarity among

Figure 2

Bayesian two-high-threshold (2HT) recognition model and clustering reveal two participant subgroups.



Note. Panel a. Posterior estimates of memory recognition (ρ) across experimental conditions. Error bars represent 89% highest posterior density intervals (HPDIs). Condition comparisons are expressed using odds ratios (ORs), with the proportion of the posterior within the region of practical equivalence (ROPE) denoted as Pr(ROPE). Panel b. Posterior correlation matrix of individual-level parameter estimates ($\boldsymbol{\beta}$), i.e., capturing effects of experimental conditions on recognition memory. Warmer colors indicate stronger positive correlations; cooler colors indicate stronger negative correlations. Matrix values represent mean posterior correlations with associated 89% HPDIs. Panel c. Individual differences in memory recognition projected onto a two-dimensional grid using Kohonen self-organizing maps (SOMs). Each node represents a prototype capturing similar observations; spatial proximity reflects similarity in the underlying feature space. Black border indicates two participant clusters identified via k-means. Accompanying individual feature maps show the distribution of each (scaled) feature across the SOM grid, illustrating multivariate relationships among parameters.

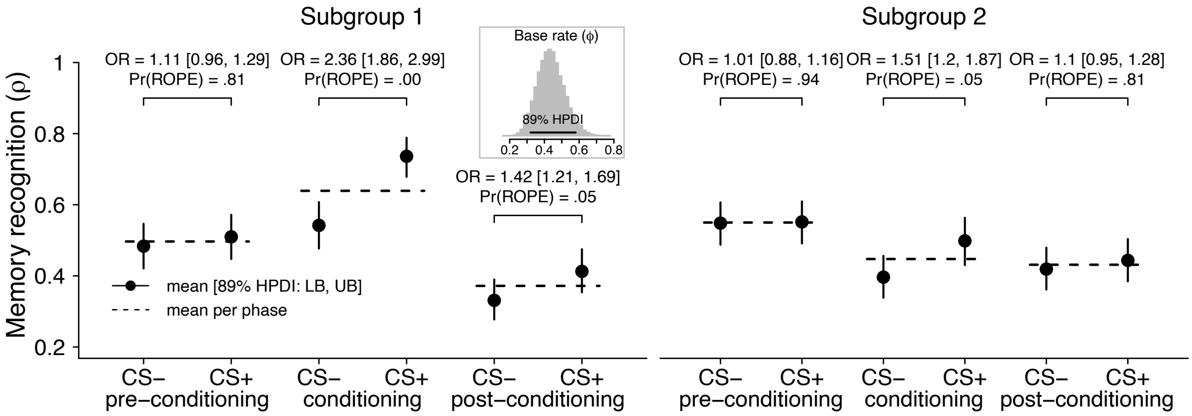
participants through spatial proximity. Specifically, maps of the grand average, phase effects, and CS+ vs. CS- differences uncovered three axes of variation: (a) a northwest-southeast gradient of category-selective prioritization (particularly during and post-conditioning); (b) a west-east gradient representing overall phase effects; and (c) a diffuse pattern of the grand average suggesting its independence.

Together, the results showed that emotional learning selectively prioritized episodic memory for items encoded during conditioning, alongside individual differences – particularly in selective prioritization and phase-related effects – that clustered into two distinct subgroups.

Emotional learning leads to two latent episodic memory profiles reflecting individual variability

Figure 3

Bayesian two-high-threshold (2HT) recognition model with latent mixture extension reveals two distinct memory profiles.



Note. Posterior estimates of memory recognition (ρ) across experimental conditions per subgroup. Error bars represent 89% highest posterior density intervals (HPDIs). Condition comparisons are expressed using odds ratios (ORs), with the proportion of the posterior within the region of practical equivalence (ROPE) denoted as Pr(ROPE). The inset plot depicts posterior proportion of the subgroups.

We formally tested for latent structure in memory recognition by extending the 2HT model to probabilistically assign participants to two subgroups, as indicated by the clustering. The estimated subgroup proportion was 0.44 [0.32, 0.58], consistent with the two-cluster solution ($\Pr_{\text{ROPE}} = .00$). These subgroups were present across all four constituent studies, suggesting they were not driven by minor variations in experimental procedures (see the Supplement). Furthermore, group membership was unrelated to individual differences in Rescorla-Wagner learning rate (r = 0 [-0.04, 0.04], $\Pr_{\text{ROPE}} = 1$) or differential anticipatory SCRs (r = -0.05 [-0.09, -0.01], $\Pr_{\text{ROPE}} = 0.97$), despite the learning rate accounting for variation in category-selective memory prioritization for items encoded during conditioning within the baseline 2HT model (see the Supplement). Both subgroups exhibited similar Pavlovian learning, with overlapping distributions of learning rates (5-number summary: Subgroup 1 = [0.01, 0.13, 0.19, 0.24, 0.53]; Subgroup 2 = [0.01, 0.12, 0.2, 0.26, 0.44]) and differential anticipatory SCRs (Subgroup 1 = [-0.26, 0, 0.02, 0.1, 0.71]; Subgroup 2 = [-0.11, 0, 0.04, 0.12, 0.78]). Overall memory recognition was comparable between the subgroups (OR = 1.13 [1.01, 1.3], $\Pr_{\text{ROPE}} = .8$), yet the detailed breakdown by condition – presented in Table 1 and depicted in Figure 3 – revealed distinct memory profiles. These results suggest that the subgroups differed in how emotional learning shaped episodic memory, rather than in associative learning performance or overall mnemonic capacity.

Profile 1: Proactive integration of memory by emotional learning

Subgroup 1 showed strong category-selective memory prioritization for items encoded during conditioning (OR = 2.36 [1.86, 2.99], $\Pr_{\text{ROPE}} = .00$), which generalized to the items encoded afterward (OR = 1.42 [1.21, 1.69], $\Pr_{\text{ROPE}} = .05$), but not beforehand (OR = 1.11 [0.96, 1.29], $\Pr_{\text{ROPE}} = .81$). When collapsing across stimulus type, recognition was highest for conditioning-phase items compared to pre- and post-conditioning phases (conditioning vs. pre-conditioning: OR = 3.39 [2.22, 5.17], $\Pr_{\text{ROPE}} = .00$; conditioning vs. post-conditioning: OR = 9.09 [6.67, 14.29], $\Pr_{\text{ROPE}} = .00$). A similar pattern was observed for CS- items alone, ruling out the possibility that the phase effect was driven solely by increased recognition of CS+ items (conditioning vs. pre-conditioning: OR = 1.27 [0.99, 1.61], $\Pr_{\text{ROPE}} = .36$; conditioning vs. post-conditioning: OR = 2.39 [1.9, 3.02], $\Pr_{\text{ROPE}} = .00$).

Table 1

Results of the two-high-threshold (2HT) model with latent mixture extension.

Subgroup	Phase	CS-	CS+	Odds ratio	Pr(ROPE)
1	pre-conditioning	.48 [.42, .55]	.51 [.45, .57]	1.11 [0.96, 1.29]	.81
	conditioning	.54 [.48, .61]	.74 [.68, .79]	2.36 [1.86, 2.99]	.00
	post-conditioning	.33 [.28, .39]	.41 [.35, .47]	1.42 [1.21, 1.69]	.05
2	pre-conditioning	.55 [.49, .61]	.55 [.49, .61]	1.01 [0.88, 1.16]	.94
	conditioning	.40 [.34, .46]	.50 [.43, .56]	1.51 [1.20, 1.87]	.05
	post-conditioning	.42 [.36, .48]	.44 [.38, .50]	1.10 [0.95, 1.28]	.81

Note. Values in CS+ and CS- columns indicate recognition probability estimates. Square brackets indicate the 89% Highest Posterior Density Interval (HPDI), representing the most credible range of parameter estimates.

Profile 2: Segmentation of memory by temporal gaps

Subgroup 2 showed an attenuated category-selective enhancement for the items encoded during conditioning (OR = 1.51 [1.2, 1.87], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .05), without generalization to pre- (OR = 1.01 [0.88, 1.16], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .94), or post-conditioning (OR = 1.1 [0.95, 1.28], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .81). This subgroup also displayed the effect of phase but in reverse: recognition was highest for items encoded before conditioning, then declined across conditioning and post-conditioning (pre-conditioning vs. conditioning: OR = 2.27 [1.67, 3], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .00; conditioning vs. post-conditioning: OR = 1.14 [0.78, 2], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .48). Again, CS- items resembled the pattern, suggesting a general phase-related decrease (pre-conditioning vs. conditioning: OR = 1.85 [1.52, 2.26], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .00; conditioning vs. post-conditioning: OR = 0.91 [0.73, 1.11], $\mathrm{Pr}_{\mathrm{ROPE}}$ = .74).

Together, two memory profiles were identified: one with strong prioritization of information encoded during and after conditioning, suggesting integration across event boundaries; and another with modest prioritization limited to conditioning, with temporal decay indicating event segmentation.

Discussion

Emotional memories are often retained because they support adaptive behavior (Nairne et al., 2007; Shohamy & Adcock, 2010). While emotional learning is known to selectively prioritize threat-related information in episodic memory, less is understood about how this influence varies across individuals – particularly for neutral events encoded close in time but in non-threatening contexts. To address this gap, we examined individual variability in the extent to which threat conditioning selectively prioritizes memory for semantically related but neutral information encoded before and after conditioning. Despite undergoing the same experimental procedure and showing similar Pavlovian learning, participants fell into two subgroups representing distinct memory profiles, suggesting individual differences in how emotional learning affects adaptive mnemonic generalization across temporal gaps.

One group showed a primacy-like pattern across the three phases, with better memory for the pre-conditioning items than for conditioning, despite the latter's emotional salience. Within conditioning, memory was selectively and modestly enhanced for items from the shock-predictive category, indicating discrimination between threat and safety. Crucially, this effect did not generalize to semantically related information from the pre- or post-conditioning phases, highlighting the role of temporal gaps in segmenting episodic memory. This profile suggests that, in some individuals, emotional learning may be constrained by event boundaries, resulting in episodic memory that predominantly reflects temporal structure.

In contrast, the other group showed a distinct pattern, with better memory for conditioning items than pre- or post-conditioning. Emotional salience appeared to disrupt the temporal decline observed in the first group. Within conditioning, memory was selectively and robustly enhanced for items from the shock-predictive category, and this prioritization extended to semantically related items encoded afterward, reflecting forward mnemonic generalization. As in the first group, no retroactive enhancement was observed, indicating asymmetric and temporally constrained memory modulation. This profile suggests that, in some individuals, emotional learning may override event boundaries, resulting in episodic memory that reflects emotional salience and integrates subsequent threat-related information encoded under safe conditions.

Importantly, our reanalysis using cognitive modeling and considering individual variability revealed no evidence of retroactive enhancement, reinforcing the finding by Kalbe and Schwabe (2021a). One explanation is that, in our dataset, pre-conditioning items were already well remembered in both groups, leaving limited opportunity for behavioral tagging – a process proposed to benefit weak memories that would otherwise be forgotten under typical conditions (Ballarini et al., 2009; Frey & Morris, 1997). In contrast, memory for post-conditioning items was weakest, and thus more likely to be selectively prioritized. Moreover, proactive strengthening may rely on distinct mechanisms, such as the carryover of emotional states that facilitate memory integration across time (Clewett & McClay, 2024; Tambini et al., 2016). Future work could test whether retroactive enhancement emerges when pre-conditioning material is weakly encoded and thus more susceptible to behavioral tagging.

The observed memory profiles may reflect how individuals perceive their environment and store threat-related experiences. In a closely related paradigm – with extinction learning replacing the neutral categorization task during post-conditioning – emotional learning selectively biased source memory, increasing misattributions of items from the shock-predictive category as having been encoded during conditioning (Laing et al., 2025). By distorting source memory, emotional learning retroactively and proactively strengthened memory for semantically related items. This raises the possibility that individuals prone to such contextual distortions may also exhibit stronger memory integration across event boundaries. However, this interpretation remains speculative, as the reported distortions reflect item-level rather than individual-level effects. In our analysis, individual differences in Pavlovian learning did not predict the observed memory profiles, suggesting that these may reflect factors beyond associative learning – perhaps such as perceived emotional salience or post-encoding consolidation (Lilja et al., 2024). These factors likely interact: lingering emotional states may blur perceived event boundaries and disrupt source memory, effectively “bridging” otherwise separate episodes, whereas effective emotion regulation may preserve normative temporal segmentation (Harris et al., 2025; McClay et al., 2023). Future work could examine whether the effects of emotional learning on episodic memory are more strongly shaped by individual differences during encoding (e.g., subjective event segmentation) or post-encoding (e.g., emotional carryover).

Limitations

Our analyses of old–new recognition judgments were implemented in the High Threshold framework, chosen for theoretical convenience rather than because it was preferable to Signal Detection Theory. Our aim was not to adjudicate between these accounts, as they typically yield qualitatively similar conclusions in standard recognition tasks. The more consequential limitation for the present study lies in the old–new task format itself (“Did you see this item?”), which is susceptible to response bias. This concern is especially relevant in our design, where new items were not uniquely tied to an encoding phase. In comparison, alternative forced-choice paradigms (“Which of these two items did you see?”) can mitigate these issues by more directly measuring relative memory strength and may provide more robust estimates in future work (Brady et al., 2022).

Conclusion

Episodic memory is naturally segmented by temporal boundaries, dividing continuous experience into discrete episodes. This segmentation creates distinct memory units and supports selective consolidation – some episodes being well retained while others fade, with emotional events typically prioritized for long-term storage. Emotional learning may enhance memory for not only the emotional episode itself but also related neutral experiences encoded outside the emotional context yet close in time. Our findings show that such cross-boundary memory enhancement varies across individuals and is most evident when the emotional episode is strongly prioritized in memory. In such cases, dominant emotional memories promote integration across later boundaries, revealing a pattern of forward mnemonic generalization. When emotional memories are less prominent, memory remains predominantly segmented according to temporal structure. Rather than uniformly enhancing memory for nearby neutral events, emotional learning reshapes episodic memory for these events, depending on how prominently the emotional episode is represented.

Declarations

General

Artificial intelligence: ChatGPT (OpenAI, Inc., US) was used for editing language.

Funding

This research was partially supported by the Alexander von Humboldt Foundation (the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research) to KM. LS received funding from the German Research Foundation, RTG2753 "Emotional Learning and Memory."

Conflicts of interest

The authors have no conflicts of interest to disclose.

Ethics approval

The original study protocol was approved by the local ethics committee at the University of Hamburg. This re-analysis of anonymized data did not require additional approval.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Open Practices Statements

Preregistration

The fourth study included in the original report was preregistered prior to data collection at https://osf.io/9hzmk. The current analysis plan was not preregistered.

Availability of data and materials

This article is based on the primary data published in Kalbe and Schwabe (2021a) and publicly available at https://osf.io/qpm3t. Demographic data and secondary data, including skin conductance response, are publicly available at [LINK]. The study materials are not publicly available due to licensing restrictions.

Code availability

Custom analysis scripts and their outputs (e.g., posterior MCMC samples and diagnostics) are publicly available at [LINK].

Author contributions

According to the CRediT taxonomy, BMB: conceptualization, data curation, formal analysis, methodology, project administration, software, validation, visualization, writing – original draft, writing – review & editing; MS: conceptualization, supervision, writing – review & editing; FK: conceptualization, writing – review & editing; KM: conceptualization, resources, supervision, writing – review & editing; LS: conceptualization, writing – review & editing.

References

Baczkowski, B. M., Haaker, J., & Schwabe, L. (2023). Inferring danger with minimal aversive experience. *Trends in Cognitive Sciences*, *27*(5), 456–467. https://doi.org/10.1016/j.tics.2023.02.005

Ballarini, F., Moncada, D., Martinez, M. C., Alen, N., & Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation. *Proceedings of the National Academy of Sciences*, *106*(34), 14599–14604. https://doi.org/10.1073/pnas.0907078106

Bartlema, A., Lee, M., Wetzels, R., & Vanpaemel, W. (2014). A bayesian hierarchical mixture approach to individual differences: Case studies in selective attention and representation in category learning. *Journal of Mathematical Psychology*, *59*, 132–150. https://doi.org/10.1016/j.jmp.2013.12.002

Betancourt, M. (2018). *A conceptual introduction to hamiltonian monte carlo*. https://arxiv.org/abs/1701.02434

Bolger, N., Zee, K. S., Rossignac-Milon, M., & Hassin, R. R. (2019). Causal processes in psychology are heterogeneous. *Journal of Experimental Psychology: General*, *148*(4), 601–618. https://doi.org/10.1037/xge0000558

Brady, T. F., Robinson, M. M., Williams, J. R., & Wixted, J. T. (2022). Measuring memory is harder than you think: How to avoid problematic measurement practices in memory research. *Psychonomic Bulletin & Review*, *30*(2), 421–449. https://doi.org/10.3758/s13423-022-02179-w

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*(4), 434–455. https://doi.org/10.1080/10618600.1998.10474787

Clewett, D., & Davachi, L. (2017). The ebb and flow of experience determines the temporal structure of memory. *Current Opinion in Behavioral Sciences*, *17*, 186–193. https://doi.org/10.1016/j.cobeha.2017.08.013

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*(3), 162–183.

https://doi.org/10.1002/hipo.23074

Clewett, D., Huang, R., & Davachi, L. (2025). Locus coeruleus activation "resets" hippocampal event representations and separates adjacent memories. *Neuron*, *113*(15), 2521–2535.e8. https://doi.org/10.1016/j.neuron.2025.05.013

Clewett, D., & McClay, M. (2024). Emotional arousal lingers in time to bind discrete episodes in memory. *Cognition and Emotion*, *39*(1), 97–116. https://doi.org/10.1080/02699931.2023.2295853

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

de Voogd, L. D., Fernández, G., & Hermans, E. J. (2016a). Awake reactivation of emotional memory traces through hippocampal–neocortical interactions. *NeuroImage*, *134*, 563–572. https://doi.org/10.1016/j.neuroimage.2016.04.026

de Voogd, L. D., Fernández, G., & Hermans, E. J. (2016b). Disentangling the roles of arousal and amygdala activation in emotional declarative memory. *Social Cognitive and Affective Neuroscience*, *11*(9), 1471–1480. https://doi.org/10.1093/scan/nsw055

Dunsmoor, J. E., Kragel, P. A., Martin, A., & LaBar, K. S. (2014). Aversive learning modulates cortical representations of object categories. *Cerebral Cortex (New York, N.Y.: 1991)*, *24*(11), 2859–2872. https://doi.org/10.1093/cercor/bht138

Dunsmoor, J. E., Kroes, M. C. W., Moscatelli, C. M., Evans, M. D., Davachi, L., & Phelps, E. A. (2018). Event segmentation protects emotional memories from competing experiences encoded close in time. *Nature Human Behaviour*, *2*(4), 291–299. https://doi.org/10.1038/s41562-018-0317-4

Dunsmoor, J. E., Martin, A., & LaBar, K. S. (2012). Role of conceptual knowledge in learning and retention of conditioned fear. *Biological Psychology*, *89*(2), 300–305. https://doi.org/10.1016/j.biopsycho.2011.11.002

Dunsmoor, J. E., Murty, V. P., Clewett, D., Phelps, E. A., & Davachi, L. (2022). Tag and capture: How salient experiences target and rescue nearby events in memory. *Trends in Cognitive Sciences*, *26*(9), 782–795. https://doi.org/10.1016/j.tics.2022.06.009

Dunsmoor, J. E., Murty, V. P., Davachi, L., & Phelps, E. A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. *Nature*, *520*(7547), 345–348. https://doi.org/10.1038/nature14106

Etz, A., Chávez de la Peña, A. F., Baroja, L., Medriano, K., & Vandekerckhove, J. (2024). The HDI + ROPE decision rule is logically incoherent but we can fix it. *Psychological Methods*. https://doi.org/10.1037/met0000660

Ezzyat, Y., & Davachi, L. (2010). What constitutes an episode in episodic memory? *Psychological Science*, *22*(2), 243–252. https://doi.org/10.1177/0956797610393742

Ezzyat, Y., & Davachi, L. (2021). Neural evidence for representational persistence within events. *The Journal of Neuroscience*, *41*(37), 7909–7920. https://doi.org/10.1523/jneurosci.0073-21.2021

Frey, U., & Morris, R. G. M. (1997). Synaptic tagging and long-term potentiation. *Nature*, *385*(6616), 533–536. https://doi.org/10.1038/385533a0

Harris, B. B., McClay, M., & Clewett, D. (2025). *Cognitive reappraisal influences the organization of emotional episodes in memory*. PsyArXiv. https://doi.org/10.31234/osf.io/kgwa3_v1

Hermans, E. J., Kanen, J. W., Tambini, A., Fernández, G., Davachi, L., & Phelps, E. A. (2016). Persistence of amygdala–hippocampal connectivity and multi-voxel correlation structures during awake rest after fear learning predicts long-term expression of fear. *Cerebral Cortex*, bhw145. https://doi.org/10.1093/cercor/bhw145

Hoffman, M. D., & Gelman, A. (2014). The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, *15*(47), 1593–1623. http://jmlr.org/papers/v15/hoffman14a.html

Kalbe, F., & Schwabe, L. (2020). Beyond arousal: Prediction error related to aversive events promotes episodic memory formation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 234–246. https://doi.org/10.1037/xlm0000728

Kalbe, F., & Schwabe, L. (2021a). On the search for a selective and retroactive strengthening of memory: Is there evidence for category-specific behavioral tagging? *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0001075

Kalbe, F., & Schwabe, L. (2021b). Prediction errors for aversive events shape long-term memory formation through a distinct neural mechanism. *Cerebral Cortex*, *32*(14), 3081–3097. https://doi.org/10.1093/cercor/bhab402

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69. https://doi.org/10.1007/bf00337288

Kruschke, J. K. (2014). *Doing bayesian data analysis: A tutorial with r, JAGS, and stan*. Academic Press.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. https://doi.org/10.1177/2515245918771304

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79. https://doi.org/10.1016/j.tics.2007.11.004

Laing, P., Hennings, A. C., Cooper, S. E., & Dunsmoor, J. E. (2025). *Emotional learning selectively distorts the temporal organization of memory: A quantitative synthesis*. PsyArXiv. https://doi.org/10.31234/osf.io/hxq3z_v1

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

Lee, M. D., & Stark, C. E. L. (2023). Bayesian modeling of the mnemonic similarity task using multinomial processing trees. *Behaviormetrika*, *50*(2), 517–539. https://doi.org/10.1007/s41237-023-00193-3

Lilja, A., Fernandez, G., & Schwabe, L. (2024). Stress enhances memory for previously encoded events depending on stressor recall. *Learning & Memory*, *31*(12), a053987. https://doi.org/10.1101/lm.053987.124

Loock, K., Heinbockel, H., Kalbe, F., & Schwabe, L. (2025). Prediction error-related memory enhancement depends on the neural state surrounding the prediction error event. *The Journal of Neuroscience*, e0739252025. https://doi.org/10.1523/jneurosci.0739-25.2025

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of effect existence and significance in the bayesian framework. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02767

McClay, M., Sachs, M. E., & Clewett, D. (2023). Dynamic emotional states shape the episodic structure of memory. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-42241-2

Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 263–273. https://doi.org/10.1037/0278-7393.33.2.263

Olsson-Collentine, A., Wicherts, J. M., & Assen, M. A. L. M. van. (2020). Heterogeneity in direct replications in psychology and its association with effect size. *Psychological Bulletin*, *146*(10), 922–940. https://doi.org/10.1037/bul0000294

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*(3), 151–160. https://doi.org/10.1037/0003-066x.43.3.151

Rouhani, N., Niv, Y., Frank, M. J., & Schwabe, L. (2023). Multiple routes to enhanced memory for emotionally relevant events. *Trends in Cognitive Sciences*, *27*(9), 867–882. https://doi.org/10.1016/j.tics.2023.06.006

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, *14*(10), 464–472. https://doi.org/10.1016/j.tics.2010.08.002

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Stan Development Team. (2023). *Stan modeling language users guide and reference manual*. Stan Development Team. https://mc-stan.org

Stan Development Team. (2024). *RStan: The R interface to Stan*. https://mc-stan.org/

Staresina, B. P., Alink, A., Kriegeskorte, N., & Henson, R. N. (2013). Awake reactivation predicts memory in humans. *Proceedings of the National Academy of Sciences*, *110*(52), 21159–21164. https://doi.org/10.1073/pnas.1311989110

Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Tambini, A., & Davachi, L. (2013). Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proceedings of the National Academy of Sciences*, *110*(48), 19591–19596. https://doi.org/10.1073/pnas.1308499110

Tambini, A., Rimmele, U., Phelps, E. A., & Davachi, L. (2016). Emotional brain states carry over and enhance future memory formation. *Nature Neuroscience*, *20*(2), 271–278. https://doi.org/10.1038/nn.4468

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., & Gelman, A. (2024). *Loo: Efficient leave-one-out cross-validation and WAIC for bayesian models*. https://mc-stan.org/loo/

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Wehrens, R., & Kruisselbrink, J. (2018). Flexible self-organizing maps in kohonen 3.0. *Journal of Statistical Software*, *87*(7), 1–18. https://doi.org/10.18637/jss.v087.i07

Williams, S. E., Ford, J. H., & Kensinger, E. A. (2022). The power of negative and positive episodic memories. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(5), 869–903. https://doi.org/10.3758/s13415-022-01013-z

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x