Emotional learning retroactively enhances item memory but distorts source attribution

Augustin C. Hennings[1,2], Jarrod A. Lewis-Peacock[1,2,3,4], Joseph E. Dunsmoor[1,2,4*]

Affiliations:

[1]Institute for Neuroscience, University of Texas at Austin.

[2]Center for Learning and Memory, Department of Neuroscience, University of Texas at Austin.

[3]Department of Psychology, University of Texas at Austin.

[4]Department of Psychiatry, Dell Medical School, University of Texas at Austin.

*Correspondence to: joseph.dunsmoor@austin.utexas.edu

Keywords: Emotion; behavioral tagging; context memory

Abstract

An adaptive memory system should prioritize information surrounding a powerful learning event that may prove useful for predicting future meaningful events. The behavioral tagging hypothesis provides a mechanistic framework to interpret how weak experiences persist as durable memories through temporal association with a strong experience. Memories are composed of multiple elements, and different mnemonic aspects of the same experience may be uniquely affected by mechanisms that retroactively modulate weakly encoded memory. Here, we investigated how emotional learning affects item and source memory for related events encoded close in time. Participants encoded trial-unique category exemplars before, during, and after Pavlovian fear conditioning. Selective retroactive enhancements in 24-hour item memory were accompanied by a bias to misattribute items to the temporal context of fear conditioning. The strength of this source memory bias correlated with participants' retroactive item memory enhancement, and source misattribution to the emotional context predicted whether items were remembered overall. In the framework of behavioral tagging: memory attribution was biased to the temporal context of the stronger event that provided the putative source of memory stabilization for the weaker event. We additionally found that fear conditioning selectively and retroactively enhanced stimulus typicality ratings for related items, and that stimulus typicality also predicted overall item memory. Collectively, these results provide new evidence that items related to emotional learning are misattributed to the temporal context of the emotional event and judged to be more representative of their semantic category. Both processes may facilitate memory retrieval for related events encoded close in time.

Introduction

2   Emotional experiences gain privileged access to the neurobehavioral mechanisms of long-term memory (Kensinger, 2009; LaBar and Cabeza, 2006; Mather et al., 2016; McGaugh, 2015;

4 Yonelinas and Ritchey, 2015). Importantly, this emotional enhancement of memory can spread to seemingly mundane details encoded close in time to the emotional experience. Through this

6 temporal association, affectively neutral information retroactively acquires the capacity to predict emotional events, allowing us to better avoid or seek out those outcomes. But what aspects of

8 memory are modulated via temporal proximity to an emotional event? Episodic memories, for instance, are composed of stimulus information (e.g., item memory) embedded with contextual

10 details (e.g., source memory) (Johnson et al., 1993; Tulving, 2002). Emotion enhances item memory (LaBar and Cabeza, 2006; Mather, 2007; Sharot and Phelps, 2004) but has inconsistent

12 effects on contextual details associated with emotional stimuli (Yonelinas and Ritchey, 2015). And while there is evidence that episodic memory is selectively prioritized for related information

14 encoded before (Dunsmoor et al., 2015) and after (Dunsmoor et al., 2015, 2018; Keller and Dunsmoor, 2020; Tambini et al., 2017) an emotional experience, retroactive and proactive effects

16 of emotional learning on source memory for this same information is unknown. Here, we investigated how temporal proximity to an emotional learning event influences both item memory

18 and contextual details for related information.

   Enhancement in memory via a temporal association between mundane and salient events

20 is consistent with neurobiological models of long-term memory. For example, the behavioral tagging hypothesis (derived from the synaptic tagging hypothesis; Frey and Morris, 1997)

22 proposes that weak learning is strengthened in memory if it is encoded within a critical time window of a more salient event and if the two events share overlapping neural ensembles

24 (Ballarini et al., 2009; Moncada and Viola, 2007; Takeuchi et al., 2016; Wang et al., 2010). As emotion is a powerful learning event, information encoded within temporal proximity may be

strengthened in memory via a mechanism of behavioral tagging. The behavioral tagging hypothesis has been translated to humans using novelty (Ballarini et al., 2013; Fenker et al., 2008; Ramirez Butavand et al., 2020), threat (Dunsmoor et al., 2015), and reward (Patil et al., 2017) to induce memory enhancements for weakly encoded information encoded close in time. However, it is unclear whether and how this mechanism may affect memory for the contextual details associated with the weak event.

Emotion can sometimes improve memory accuracy for information directly associated with an emotional stimulus, such as the spatial and temporal context in which the emotional stimulus was encoded (Kensinger and Schacter, 2006; Rimmele et al., 2012; Schmidt et al., 2011; Talmi et al., 2019). Therefore, one possibility is that emotional learning improves both item and source memory for information in temporal proximity to an emotional event. In this case, we might expect item memory to be accompanied by source memory accuracy, such that items are appropriately organized in time relative to the emotional event. Alternatively, emotional learning might have no effect (Sharot and Yonelinas, 2008; Wang and Fu, 2010) or even impair memory for contextual information. Indeed, it may be simpler to remember a trivial item by virtue of associating the item with the salient context (Takashima et al., 2016). It is therefore plausible that linking weak and strong events by temporal proximity might improve item memory at the cost of source memory. Consequently, we may expect a relationship between the strength of item memory for related events encoded close in time and the strength of a misallocation bias to source the item to the more salient temporal context. This might suggest that retroactive memory enhancement for weakly encoded items relies in part on misattribution to the more salient emotional context.

In addition to our exploration of how emotional learning impacts item and source memory for proximal events, we tested a parallel hypothesis that emotional learning alters abstract stimulus properties that may indirectly facilitate memory retrieval for those items as well. Previous work demonstrates that emotional learning is sensitive to how well an item represents its broader

category, i.e., typicality (Dunsmoor and Murphy, 2014, 2015; Dunsmoor et al., 2014; Lei et al., 2019; Struyf et al., 2018). Whether emotional learning has the power to alter an abstract stimulus property such as subjective typicality is unknown, but stimulus memorability may be an underappreciated factor contributing to item memory (Bainbridge, 2019). For instance, the hippocampus plays a role in both episodic memory and concept representations (Davis and Poldrack, 2014; Mack et al., 2016; Quiroga, 2012). Moreover, typical category members used as conditioned stimuli in Pavlovian fear conditioning preferentially engages the hippocampus and hippocampal-amygdala functional connectivity (Dunsmoor et al., 2014), which is a substrate for emotional memory enhancement (Murty et al., 2010).

In the present study participants underwent a two-day Pavlovian fear conditioning task that included trial-unique (i.e., non-repeating) pictures of animals and tools as conditioned stimuli (CSs), based on the protocol from Dunsmoor et al. (2015). Items were encoded before, during, and after Pavlovian fear conditioning. We predicted that emotional learning would have divergent effects on 24-hour episodic memory accuracy for CS items and the encoding temporal context. Specifically, we predicted that CSs semantically related to the fear conditioned category would be selectively remembered regardless of their temporal context (Dunsmoor et al., 2015), but that participants would have a bias to attribute these related CSs to the temporal context of fear conditioning. We also investigated whether emotional learning enhances subjective typicality for CS category members related to the fear conditioning category, and whether long-term memory would be influenced by how strongly an item was deemed to represent its superordinate category. Such findings might indicate that weak memories formed in temporal proximity to an emotional experience our bound to the more salient temporal context, and that emotional learning can alter abstract stimulus properties of weakly encoded information to make this information more memorable.

Results

During fear conditioning, participants learned through experience that exemplars from one category (CS+, animals or tools, counterbalanced between participants) were paired with an aversive electrical shock unconditioned stimulus (US) to fingers of the right hand (50% CS-US pairing rate). Exemplars from the other category (CS-, tools or animals, respectively) were never paired with shock. Each encoding phase was separated by a short (<5 minute) break. The following day (~24 hours later), participants underwent a recognition memory test that included each 'old' exemplar plus 'new' pictures of animals and tools. The recognition memory test was then followed by a separate source memory test that asked participants to place each item encoded the previous day into the context in which it was presented (3 alternative-forced-choice, AFC). Participants also rated the typicality for each basic-level exemplar in relation to its superordinate category (animal, tool). See *Materials and Methods* for further details on the experimental design.

24-hour recognition memory

In order to determine the effect of an emotional event on memory for temporally related items, we collected 24-hour delayed recognition memory for items encoded before, during and after fear conditioning. Analysis of recognition memory performance revealed better memory for items from the CS+ compared to the CS- category for all temporal contexts (**Fig. 1**). Memory performance was calculated as corrected recognition (high confidence hits minus high confidence false alarms). The false alarm rate did not differ between CS+ and CS- (2-sided Wilcoxon signed-rank test P = 0.13, CLES = 0.55; see **Supplementary Tables 1 & 2** for full recognition memory results). A repeated measures ANOVA with CS Type (CS+, CS-) and Temporal Context (pre-conditioning, conditioning, post-conditioning) revealed a main effect of CS ($F_{1, 44}$ = 28.66, P_{perm} = 0.0001, η^2_G = 0.073) and Context ($F_{2, 88}$ = 19.72, P_{perm} = .0001, η^2_G = 0.092), as well as a CS by Context interaction ($F_{2, 88}$ = 9.04, P_{perm} = 0.0002, η^2_G = 0.020). Follow-up planned comparisons showed enhanced memory for CS+ items encoded during fear conditioning (2-sided Wilcoxon

signed-rank test P = 5e-6, CLES = 0.72). This selective enhancement in memory was retroactive

for CS+ compared to CS- items encoded during pre-conditioning (P = .0055, CLES = 0.59), and

proactive for CS+ compared to CS- items encoded post-conditioning (P = .0053, CLES = 0.61).

Thus, recognition memory results replicate previous findings that Pavlovian conditioning

enhances memory for CS+ trials encoded before, during, and after fear conditioning (Dunsmoor

et al., 2015).
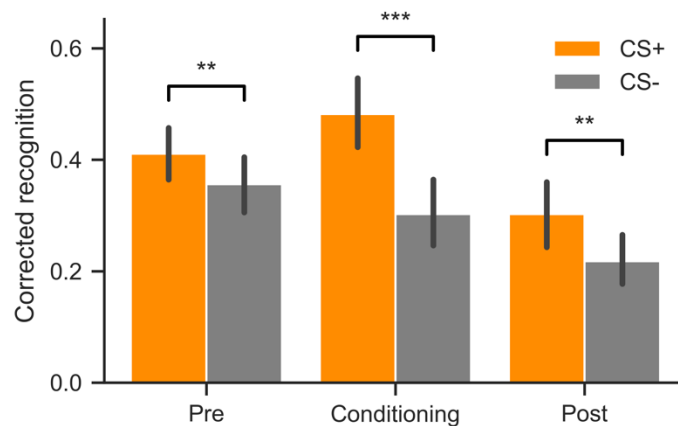
Fig 1. Emotional learning retroactively and proactively enhances recognition memory for related information. CS+ items were remembered more than CS- for all temporal contexts. Corrected recognition was calculated as high confidence hits minus high confidence false alarms. Error bars indicate the 95% confidence interval around the mean. In all experimental phases recognition memory was higher for CS+ items compared to CS- items. **P < 0.01, ***P < 0.001

Temporal context memory

In order to compare the effects of an emotional event on item vs. context memory, we collected

source temporal context judgements for each item. Source memory judgements for the temporal

context (pre-conditioning, fear conditioning, post-conditioning) associated with each CS from

encoding were calculated as the proportion of items per CS type attributed to each context as a

function of when the item was encoded (**Fig 2**). A 3-way repeated measures ANOVA with CS

Type, Temporal Context, and Source Memory Response revealed a significant main effect of

2   Source Memory Response ($F_{2, 66} = 10.24$, $P_{perm} = 0.0004$, $\eta^2_G = 0.12$), an interaction of CS Type

and Source Memory Response ($F_{2, 66} = 17.87$, $P_{perm} = 0.0001$, $\eta^2_G = 0.13$), an interaction of

4   Temporal Context and Source Memory Response ($F_{4, 132} = 6.09$, $P_{perm} = 0.0005$, $\eta^2_G = 0.027$), and

a significant 3-way interaction ($F_{4, 132} = 3.15$, $P_{perm} = 0.0154$, $\eta^2_G = 0.013$). Only CS+ items encoded

6   during fear conditioning and CS- items encoded during post-conditioning (i.e., extinction) were,

on average, attributed to their appropriate temporal context. Significance was determined by

8   testing the mean response proportion against chance (33%) using a bootstrap resampling

procedure (Davison and Hinkley, 1999). There was a bias to misattribute CS+ items from each

10   encoding phase to the fear conditioning context regardless of veracity (all mean proportions

greater than chance; see **Supplementary Table 3** for full temporal context memory results). This

12   bias was selective to the CS+ vs. CS- for pre-conditioning (2-sided Wilcoxon signed-rank test P

= 7.07e-4, CLES = 0.68), fear conditioning (P = 1.21e-4, CLES = 0.80) and post-conditioning (P

14   = 0.0017, CLES = 0.74).
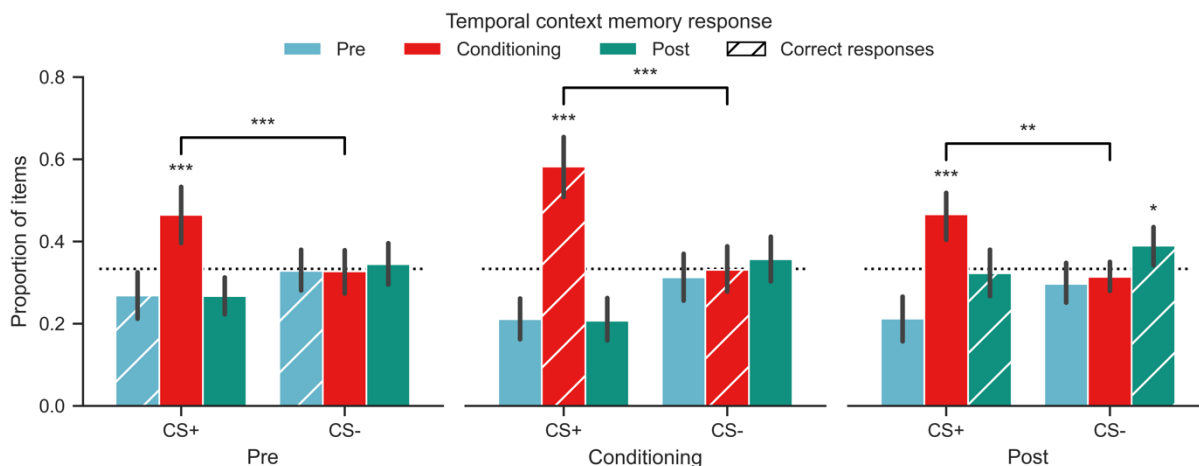
16   **Fig 2. Emotional bias in source memory attribution**. CS+ items
from all temporal contexts were endorsed as being from the fear
18   conditioning temporal context more than CS- items, regardless of
when the items were actually encoded. Responses are shown as
20   the proportion of total items for each CS type encoded in each
temporal context (1.0 = 24 items). Error bars correspond to the
22   bootstrapped 95% confidence interval. Hatched bars indicate the

Association between selective memory enhancements and source memory bias

For each phase of encoding, there was a bias to recognize more CS+ than CS- items as well as a bias to report CS+ items as having been encoded during fear conditioning regardless of when they were encountered. One possibility is that there is an association between these biases in item and source memory. Specifically, participants who exhibit a strong memory selectivity for CS+ items encoded before, during, and after fear conditioning may also be biased in attributing these items to the fear conditioning temporal context. To explore this possibility, we examined the correlation between each participants recognition memory bias (i.e. CS+ minus CS- corrected recognition from each phase; data shown in **Fig 1.**) against the fear conditioning temporal context bias (i.e. the bias to attribute CS+ minus CS- items to the temporal context of fear conditioning; data shown in **Fig 2.**). We observed a significant robust Pearson's correlation for pre-conditioning ($r_{skipped} = 0.34$, 95% CI = [0.07, 0.59], $P = 0.0188$), fear conditioning ($r_{skipped} = 0.36$, 95% CI = [0.06, 0.61], $P = 0.0194$), and post-conditioning ($r_{skipped} = 0.37$, 95% CI = [0.03, 0.66], $P = 0.0304$). This suggests that selective recognition memory for CS+ items at each stage of encoding, including before and after fear conditioning, was related to the tendency to selectively misattribute CS+ items to the more salient temporal context.
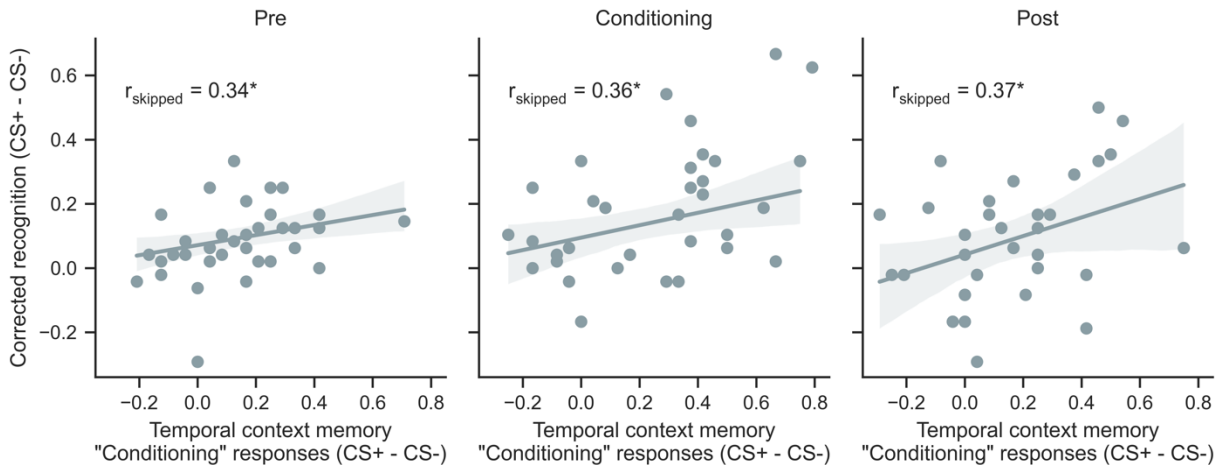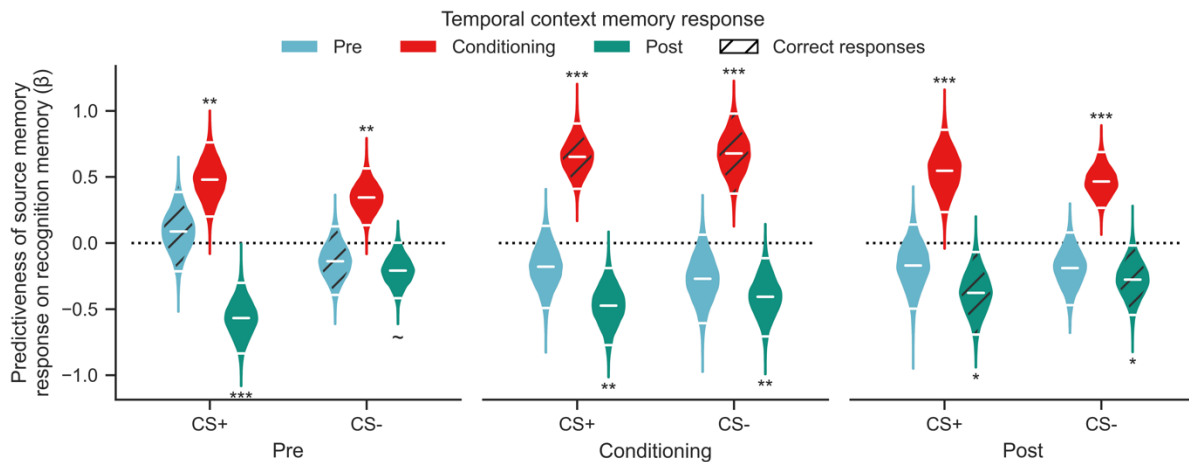
Fig 3. Temporal context memory bias predicts recognition memory enhancement. CS+ - CS- difference in fear conditioning source memory responses were correlated with CS+ - CS- differences in corrected recognition for each phase. Robust correlation was used, eliminating the effects of outliers. All data points are shown. Shaded region corresponds to the 95% CI of the skipped Pearson's correlation. *$P < 0.05$

Item recognition is influenced by temporal context judgements

To further test our hypothesis that neutral items are remembered better if they are misattributed to the emotional context, we conducted a multiple logistic regression to predict recognition memory as a function of the item's source memory judgement (**Fig. 4**). Each regression was run as a fixed-effect analysis across all participants, and a bootstrap resampling procedure was used to test for generalizability. The significance of the logistic regression coefficients was determined by testing the obtained bootstrap distribution against zero. Results showed that for all CS+ and CS- items encoded before, during, and after fear conditioning, attributing items to the fear conditioning context was positively linked to accurate recognition memory (All "Conditioning" coefficients greater than 0, see **Supplementary Table 4** for full results). In contrast, attributing CS+ or CS- items to the post-conditioning (extinction) context was negatively linked to recognition memory, indicating items were more likely to be forgotten. This

pattern of attributing forgotten items to the extinction context was observed for items encoded in each phase. These results show that remembered items were more likely to be judged as having been encoded during fear conditioning, regardless of CS type and regardless of when they were encoded.


Fig 4. Source memory misattribution to salient context predicts recognition memory. For each CS type in each temporal context, source memory responses were used to predict recognition memory high confidence hits. Violins show the obtained distribution of logistic regression coefficients, horizontal white lines correspond to the mean and 95% CI. Each triplet of violins corresponds to a single logistic regression. Significance markers indicate if the mean of the distribution is reliably different from zero. Hatched violins are responses that were correct on the 3-AFC temporal context memory test. ˜$P = 0.055$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Emotional learning alters perceived typicality

At the conclusion of the experiment on Day 2, participants rated the typicality of each item encoded the previous day to determine how representative each exemplar was to its category, animal or tool (**Fig. 5A**). A CS Type by Temporal Context repeated measures ANOVA revealed a main effect of CS ($F_{1,33} = 6.25$, $P = 0.018$, $\eta^2_G = 0.043$). There was no main effect of Temporal Context, nor was there a CS by Context interaction (Ps > .72). Follow up comparison revealed that CS+s were rated as more typical than CS-s for pre-conditioning (2-sided paired $t_{33} = 2.12$, P

= 0.041, d = 0.36), fear conditioning (t_{33} = 2.22, *P* = 0.033, d = 0.43), and post-conditioning (t_{33} =

2.40, *P* = 0.022, d = 0.47). These data suggest that fear conditioning alters the perceived typicality

of basic-level category exemplars to their superordinate category. See **Supplementary Fig. 1** for

a stimulus-wise comparison of the difference in typicality between CS+ and CS-. The category

(animal or tool) that was used as the CS+ during fear conditioning was counter balanced across

participants. One possibility is that the typicality effect was influenced by which category was

paired with shock. In order to rule out an effect of which category was used as the CS+, a repeated

measures ANOVA was run with CS type as a within-subjects factor, and shocked-category as a

between-subjects factor and. There was no significant main effect or interaction of shocked-

category (*P*s > .15), thus the effect of emotional learning on perceived typicality was not different

in animals vs. tools. These data provide new evidence that fear conditioning alters the perceived

typicality of category exemplars associated with an aversive outcome.
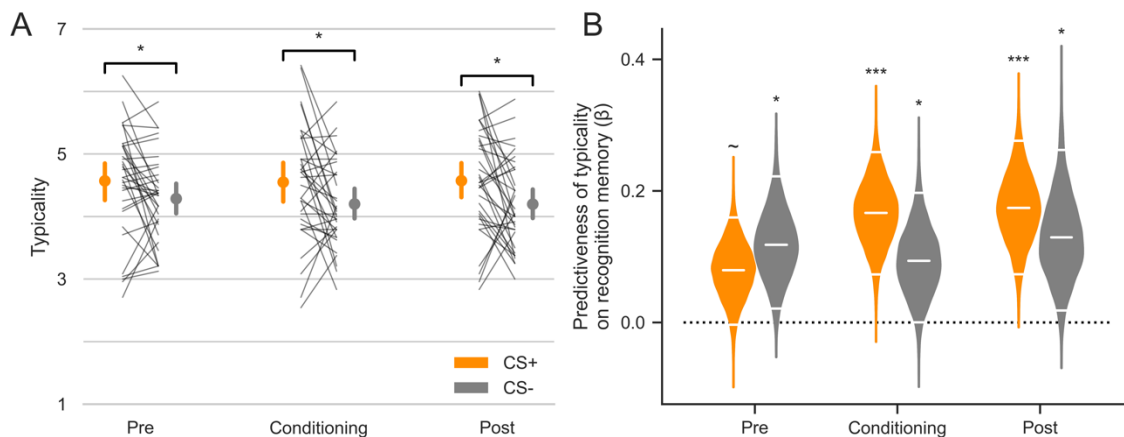
Fig 5. A. Aversive learning increases perceived typicality. CS+ items from all temporal contexts were rated as more typical than CS- items. Items were scored from 1 (atypical) to 7 (very typical). CS+ category (animals or tools) was counterbalanced across participants. The group mean and 95% CI are displayed next to individual participants data. **B. Perceived typicality predicts recognition memory**. Violins show the obtained distribution of logistic regression coefficients linking typicality ratings and item recognition. Horizontal white lines correspond to the mean and 95% CI. Each violin corresponds to a single logistic regression.

Typical is memorable

Given our secondary hypothesis regarding a potential role of typicality on selective episodic memory enhancements, we assessed if perceived stimulus typicality was related to item recognition memory. For each CS type in each temporal context, a logistic regression was run using typicality ratings (1-7) to predict high confidence recognition memory hits (**Fig 5B**). Typicality was predictive of recognition memory for both CS+ and CS- in all temporal contexts (all Ps < 0.057, see **Supplementary table 5** for full results). There was no difference in the strength of this relationship between CS+ and CS- in any phase (Ps > .30). These results indicate that items that were perceived as more typical were also more likely to be remembered, but this relationship was not affected by CS type nor encoding temporal context.

Relative contributions of source memory misattribution and typicality on memory.

Our results indicate that fear conditioning enhances recognition memory for related items (**Fig 1.**). This emotional learning also results in an increase in source memory misattribution to the conditioning context for CS+s (**Fig. 2**), and an increase in perceived typicality for CS+s (**Fig. 5A**). These source memory misattributions and typicality ratings are in turn also predictive of recognition memory (**Figs. 4 & 5B**). We then assessed the relative strength of all three factors (CS type, conditioning source context misattributions, and typicality) on recognition memory. This is of particular interest for items encoded during pre-conditioning, where these effects emerge retroactively after emotional learning. Multiple logistic regressions were run for each participant that included CS type and conditioning source memory misattributions as binary predictors, and typicality ratings as a continuous predictor of recognition memory for items encoded during pre-conditioning. Unsurprisingly given our previous results, all three factors were positively linked to

recognition memory (two-tailed t-test against 0, CS type: $t_{33} = 3.40$, $P = 0.002$, d = 0.58; Conditioning source misattributions: $t_{33} = 5.03$, $P = 1.7e-5$, d = 0.86; Typicality: $t_{33} = 2.92$, $P = 0.006$, d = 0.50). Source memory misattribution to conditioning was a more powerful predictor of recognition memory than both CS type (paired two-tailed $t_{33} = 2.21$, $P = 0.034$, d = .46) and typicality ($t_{33} = 3.32$, $P = 0.002$, d = 0.73). There was no difference between CS type and typicality ($t_{33} = 1.06$, $P = 0.30$, d = 0.26). These results indicate that source memory misattributions to the conditioning context are the most predictive factor of the retroactive enhancement in recognition memory observed in this study.
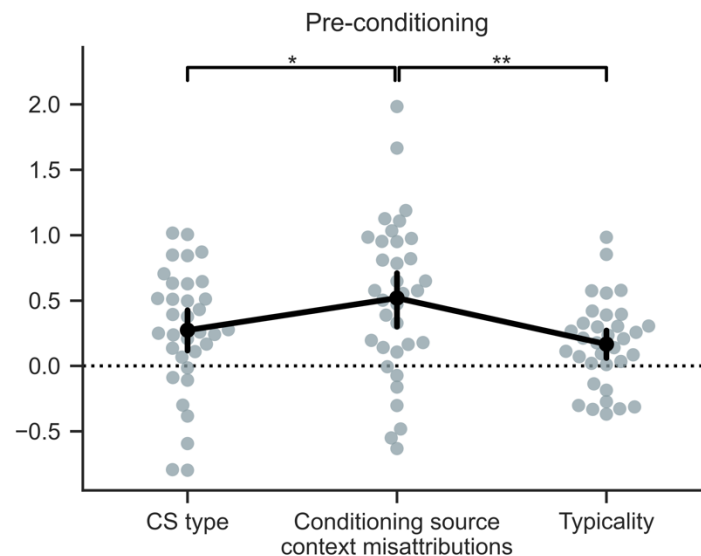
Fig 6. Source context misattribution is the most predictive feature of retroactive recognition memory enhancement. All three factors were significantly predictive of recognition memory (above 0). There was no difference between CS type and typicality. Black points and bars indicate the group mean and 95% CI. *P < 0.05, **P < 0.01.

Discussion

While emotional events are often prioritized in memory, it is far less clear how emotional events affect memory for other information encoded close in time. Here, we provide support for a

behavioral tagging mechanism in human episodic memory, whereby emotional learning

selectively and retroactively enhanced item memory conceptually related to the emotional event,

replicating prior findings (Dunsmoor et al., 2015; Patil et al., 2017). Building on these prior findings,

we found that related information encoded before or after fear conditioning is consistently

misattributed to the more salient fear conditioning temporal context. Put in terms of a tag-and-

capture mechanism (Redondo and Morris, 2011): memory attribution was biased to the temporal

context of the strong event that provided the source of memory stabilization for the weak event.

We also found new evidence that emotional learning selectively enhances subjective typicality for

related exemplars, and that typical items were better remembered overall. Collectively, these

findings advance knowledge on how emotional learning affects mnemonic and abstract stimulus

representations for information encoded before and after an emotional experience.

While effects of emotional arousal on item memory are fairly consistent (LaBar and

Cabeza, 2006; McGaugh, 2015; Murty et al., 2010), the effect of emotion on memory for

contextual details is less straightforward. Contextual details sometimes receive a boost in memory

(Rimmele et al., 2012; Schmidt et al., 2011), but other studies find no effect or worse memory for

contextual details associated with an emotional item (Sharot and Yonelinas, 2008; Takashima et

al., 2016; Wang and Fu, 2010). The divergent effects of emotion on item and source memory may

be related to different temporal dynamics in forgetting between item-emotion binding, supported

by the amygdala, and item-context binding, supported by the hippocampus (Yonelinas and

Ritchey, 2015). Specifically, amygdala-dependent emotional item memory is resistant to

forgetting, whereas hippocampus-dependent item-context memory decays at a faster rate.

Notably, much of the work on the role of emotional arousal on item versus source memory uses

memoranda that has an intrinsically emotionally arousing feature (e.g., evocative images with a

contextual detail; but see Wang and Fu, 2010). Extending this model to *neutral* memoranda

encoded *around the time* of emotional learning has interesting implications for interpreting

emotional enhancements of episodic memory through the lens of a putative behavioral tagging

2 mechanism. Specifically, it suggests that emotional events act predominately on the amygdala to

support retroactive and proactive item memory enhancements for related information. An

4 amygdala-based item memory enhancement would accord with the use of the Pavlovian

conditioning design employed here. In this way, amygdala-dependent fear conditioning may

6 upregulate hippocampal processing to support selective consolidation of otherwise mundane

items encoded close in time. In an emotional-binding framework (Yonelinas and Ritchey, 2015),

8 this retroactive item-emotion memory benefit would not necessarily be accompanied by a

retroactive hippocampus-dependent item-context memory benefit for those same weakly

10 encoded items.

One important question is the relative contribution of consolidation versus retrieval

12 processes on the bias to misattribute CS+ items to the temporal context of fear conditioning. That

is, is item memory enhanced because the item is bound to the emotional context, or is the item

14 attributed to the emotional context because it is remembered? One possibility is that emotional

learning boosted memory for weakly encoded items by linking these memories to the temporal

16 context that provided the source of memory stabilization (Clewett et al., 2020). This explanation

is in keeping with a tag-and-capture model, which is primarily a mechanistic account for the

18 consolidation of newly formed memories (Redondo and Morris, 2011). This account may afford

some explanatory power for the strong correlation between selective item memory enhancements

20 (CS+ > CS-), which our prior work shows is consolidation dependent (Dunsmoor et al., 2015,

2018; Patil et al., 2017), and the bias to misattribute CS+ items to the fear conditioning context.

22 Alternatively, source memory errors may be a factor of retrieval processes, per se. One

possibility is that memory retrieval of a CS+ exemplar may be accompanied by memory retrieval

24 of the salient fear conditioning context. As a consequence, participants may have a bias to

attribute more CS+ items to the emotional context because those items are accompanied by a

stronger feeling of remembering (Sharot et al., 2004) induced by reactivation of the emotional

context. This explanation is consistent with the general properties of source monitoring that involve *retrospectively* attributing the origin of a memory to a particular source (Johnson et al., 1993), and with mental context reinstatement models of memory retrieval (Gershman et al., 2013; Howard and Kahana, 2002a; Smith and Vela, 2001; Tulving and Thomson, 1973). It is also consistent with a retrieval focused account of emotional memory enhancements (Talmi et al., 2019), wherein emotional item memory is bound to changing representations of mental contexts. Interestingly, there is evidence that distinct phases of a fear conditioning experiment can be understood as separable mental contexts that uniquely organize memory (Dunsmoor et al., 2018; Hennings et al., 2020) in a manner consistent with how shifts in mental context segment experience for distinct episodes (Clewett et al., 2019; DuBrow et al., 2017). That source memory misattributions were more strongly predictive of the retroactive enhancement in item memory compared to CS type (**Fig. 6**) may also support the importance of a retrieval mechanism. Put together, it is possible that remembering a CS+ item reactivates the mental context associated with the fear conditioning phase, thereby promoting retrieval for related items, including those encoded in temporal proximity before or after the fear conditioning phase.

We also found that items from the CS+ category were rated as more typical than CS-items, regardless of when the item was encoded. One constraint to this experimental design is the inability to gather typicality ratings from each participant prior to the encoding session, as pre-exposing items prior to encoding would necessarily interfere with interpretation of the episodic memory results. But importantly we showed that there was no inherent difference in typicality as a function of superordinate category (animal, tool), suggesting that fear conditioning itself modulated typicality judgements. These results are new evidence that conditioning can alter an abstract stimulus property like how well an item represents its category. This enhancement in representativeness may enable CS+ items to be better remembered over time, as we also found

that typical items were more memorable overall. There is surprisingly little research on the link between stimulus typicality and memorability. However, there is evidence that semantic similarity promotes free recall (Howard and Kahana, 2002b), and abundant evidence that the hippocampus plays a role in concept representation and categorization (Davis and Poldrack, 2014; Mack et al., 2016; Quiroga, 2012). Thus, one mechanistic possibility is that fear conditioning augmented semantic organization by raising the overall typicality of distinct category members from the CS+ category, which in turn facilitated memory for related items encoded close in time. More research is warranted to help elucidate the relationship between typicality and memorability more broadly.

Recent work has made progress on translating neurobiological models of long-term memory to explain memory strengthening for weak memories encoded in temporal proximity to a more salient event. According to the synaptic tagging hypothesis (Frey and Morris, 1997; Redondo and Morris, 2011), and its behavioral counterpart (Moncada and Viola, 2007), weak activation that is only sufficient to produce a transient short-term memory can be transformed into a durable long-term memory if accompanied by a more salient event that relies on the same neural system within a critical time window. The weak event sets a putative learning tag that is stabilized by release of plasticity related proteins induced by the strong event, broadening the window of time by which salient events can modulate long-term memory. But memories are multifaceted, and how a putative behavioral tagging mechanism impacts separate mnemonic aspects of the same experience, or other abstract stimulus properties, is not well understood. Matching the task demands of episodic memory and Pavlovian conditioning allowed us to show that an emotional event has divergent effects on item and source memory and enhances subjective typicality of related events encoded close in time. Mechanisms that link seemingly inconsequential information to a future or past emotional event can serve an adaptive function. For instance, we do not always know the significance of numerous details we encounter throughout the day. By integrating memories encoded close in time to a meaningful event, we can

remember information that may be relevant to seeking out or avoiding similar outcomes in the future. Enhancing the representativeness of the experience may serve a similar mnemonic function. But linking memories by temporal proximity may come at the cost of misattributing non-emotional memories to an emotional context and over interpreting the representativeness of information related to an emotional event. Such an organization proves maladaptive when innocuous cues experienced around the time of highly negative events trigger retrieval of unwanted emotional memories. Thus, as a consequence of organizing memories based on temporal proximity to a salient event, a host of potentially irrelevant and innocuous information acquires the capacity to reactivate emotional memories in the future.

Materials and Methods

Participants. This experiment was approved by the University of Texas at Austin IRB (IRB # 2017-02-0094). A total of 48 adult volunteers (mean age 24 y/o, s.d. 4.5 years; 32 female) were recruited and completed a two-day functional MRI investigation. An additional 3 participants were recruited but did not complete the study. A primary aim of the neuroimaging component of the investigation was to examine the neural correlates of fear extinction memory encoding and retrieval. As such, fMRI and behavioral data pertaining to fear conditioning and extinction are reported in Hennings et al. (2020). We recruited participants who were psychologically healthy (N = 24) as well as participants who reported PTSD symptoms (N = 24). For the purposes of the

memory analyses, and given no *a priori* hypotheses regarding episodic memory results between these groups, we combined participants for this report. We confirmed highly similar results in both groups; specifically, in the analysis of 24-hour recognition memory (**Fig. 1**) a mixed ANOVA revealed no main effect or interaction with Group (Healthy vs. PTSD symptoms). 3 additional participants were removed from analysis due to extremely poor memory performance that was near chance (N = 1), or a failure to make a response during the recognition memory test on > 10% of trials (N = 2). Finally, 11 participants were unable to complete the source memory test due to time constraints at the MRI facility, and hence data from these participants were not included in the temporal context memory and typicality results. After exclusions, our sample was thus N = 45 participants for the analysis of recognition memory, and N = 34 participants for the analyses of temporal context memory and typicality.

Conditioned and unconditioned stimulus. Memoranda (i.e., CSs) included color photographs of animals and tools presented on a white background, obtained from the website http://www.lifeonwhite.com or publicly available resources on the internet. Each CS was a distinct basic-level exemplar with a unique name; in other words, there were not two different pictures of a bear throughout the entire experiment. The order of stimuli was counterbalanced and pseudorandomized such that no more than three pictures from the same category occurred in a row. Stimulus presentation was controlled using E-Prime 3 (https://pstnet.com/products/e-prime/). The US was a 50-millisecond electrical shock delivered to the index and middle fingers of the left hand. The US was calibrated prior to the start of the experiment to reach a level deemed "highly annoying but not painful." The shock was controlled using the STMEPM-MRI stimulation system from BIOPAC Systems (Goleta, CA).

Psychophysiology.

Task design. The paradigm was based in large part on Dunsmoor et al., (2015). Day 1 included pre-conditioning, fear conditioning, and post-conditioning (i.e., extinction). Each phase included

24 CS+ and 24 CS- trials, for a total of 144 trials on Day 1. A short (~1 minute) break was interspersed between each encoding phase. During pre-conditioning, participants viewed pictures of animals and tools and made a 2-AFC category judgement (animal, tool) using a button box in their right hand. During fear-conditioning, 50% of the images from one category (CS+) trials co-terminated with the US, and participants rated shock expectancy (Yes/No). CS+ category was counterbalanced across participants. During post-conditioning, participants continued to rate shock expectancy but no shocks were delivered (i.e., extinction). CS duration in all phases on Day 1 was 4.5 ± .05 s (jittered) followed by an intertrial interval of 6 ± 1s. During pre-conditioning and fear conditioning, the intertrial interval was a blank background with a fixation cross. For reasons detailed in Hennings et al. (2020), post-conditioning (i.e., extinction) included a rapidly presented series of scene pictures (5, 6, or 7 pictures presented for 1 s each) during the intertrial interval. The rationale for presenting scene pictures during the ITI was for multivariate pattern analysis of the extinction context, but is not related to the purposes of the episodic memory data reported here. Participants returned ~24 hours later and underwent a test of fear renewal (data reported in Hennings et al., 2020), and a surprise recognition memory test inside the scanner. The recognition memory test included 144 old pictures from the previous day and 96 new pictures of animals and tools. Participants had 3 seconds to rate each item as old or new and their level of confidence (definitely old, maybe old, maybe new, definitely new). Trial order was pseudorandomized to ensure a balance of old and new stimuli from each category from each encoding phase.

Finally, participants completed the source memory test outside the scanner in a different room. The source memory test included only old images from Day 1, and participants were informed that each item they saw had been seen the previous day. Participants were reminded that the experiment had been divided into three different phases (labeled Phase 1, 2 and 3), and were given a brief reminder of each phase (i.e. "Phase 1 was at the beginning when you were

classifying each image as an animal or a tool"). Participants were instructed to respond to their

2 best of the ability when they had seen each picture.

After the source memory instructions, participants were instructed on how to rate the

4 typicality of each image on a 7-point scale (1= Not at all typical; 7 = Very typical). An example of

typicality was given using an unrelated category: "For example, an apple is a typical fruit, it is

6 representative of the entire fruit category. On the other hand a dragonfruit is not a typical fruit, at

least in American grocery stores." The task was self-paced, with a brief 500ms or 800ms,

8 randomly chosen, ITI after participants had answered both questions for each picture.

Statistical analyses.

10 Behavioral data was analyzed using a combination of Python and R. Corrected recognition

and temporal context memory were determined to be non-normally distributed (Shapiro-Wilk test

12 Ps <.05), and thus were analyzed with non-parametric statistics (Shapiro and Wilk, 1965). When

mentioned, the number of permutations or bootstrap iterations was set to 10,000, which allows

14 for a minimum empirical p-value of 0.0001. All data were first submitted to an ANOVA using the

R package *ez* (Lawrence, 2016), and the package *permuco* was used to permute ANOVAs for

16 non-normal data (Frossard and Renaud, 2019). ANOVA effect sizes are reported as generalized

eta squared (Bakeman, 2005), the parametric effect size is reported for non-normal data for

18 clarity. Student's t-tests or Wilcoxon sign-ranked tests were used for follow up comparisons using

the Python package *pingouin* (Vallat, 2018). Effect size for t-test is reported as Cohen's d, and

20 the effect size for Wilcoxon sign-ranked tests is the common language effect size (CLES; Vargha

et al., 2000).

22 For the analysis correlating the emotional bias in temporal context memory to the

emotional bias in recognition memory, a robust skipped-Pearson's correlation was used for non-

24 normal data (Pernet et al., 2013). After the removal of any outliers, the data were resampled with

replacement and significance was determined by testing the obtained distribution of Pearson's r

2   values against 0. Logistic regressions analyses were performed using the Python package *scikit-*

learn (Pedregosa et al., 2012) with the lbfgs solver and L2 regularization. For the regression using

4   source memory to predict recognition memory hits, 3-AFC temporal context memory responses

each CS type in each temporal context the were transformed into a one-hot encoding scheme.

6   That is, binary coding across three columns (one for each possible 3-AFC response), was used

to predict high-confidence recognition memory hits. In order to overcome a limited number of

8   certain trial types in some participants (e.g. a participant with few or no pre-conditioning CS+

items accompanied by a source memory attribution to the "post-conditioning" phase), each logistic

10   regression was run as a fixed-effect analysis across all participants. To test for generalizability,

we implemented a bootstrapping procedure in which whole participants were sampled with

12   replacement on each iteration. Significance of regression coefficients was determined by testing

the obtained distribution against 0. The same fixed-effects bootstrap procedure was used in the

14   regression using typicality to predict recognition memory.

References

Bainbridge, W.A. (2019). Memorability: How what we see influences what we remember. In Psychology of Learning and Motivation, K.D. Federmeier, and D.M. Beck, eds. (Academic Press), pp. 1–27.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. Behav. Res. Methods *37*, 379–384.

Ballarini, F., Moncada, D., Martinez, M.C., Alen, N., and Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation.

Ballarini, F., Martínez, M.C., Díaz Perez, M., Moncada, D., and Viola, H. (2013). Memory in Elementary School Children Is Improved by an Unrelated Novel Experience. PLoS One *8*, 66875.

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

Clewett, D., Dunsmoor, J.E., Bachman, S., Phelps, E., and Davachi, L. (2020). Survival of the salient: Emotion rescues otherwise forgettable memories via neural reactivation and post-encoding hippocampal connectivity. BioRxiv 2020.07.07.192252.

Davis, T., and Poldrack, R.A. (2014). Quantifying the internal structure of categories using a neural typicality measure. Cereb. Cortex *24*, 1720–1737.

Davison, A., and Hinkley, D. (1999). Bootstrap methods and their application (Cambridge New York: Cambridge University Press).

DuBrow, S., Rouhani, N., Niv, Y., and Norman, K.A. (2017). Does mental context drift or shift? Curr. Opin. Behav. Sci. *17*, 141–146.

Dunsmoor, J.E., and Murphy, G.L. (2014). Stimulus Typicality Determines How Broadly Fear Is Generalized. Psychol. Sci. *25*, 1816–1821.

Dunsmoor, J.E., and Murphy, G.L. (2015). Categories, concepts, and conditioning: How humans generalize fear. Trends Cogn. Sci. *19*, 73–77.

Dunsmoor, J.E., Kragel, P.A., Martin, A., and La Bar, K.S. (2014). Aversive learning modulates cortical representations of object categories. Cereb. Cortex *24*, 2859–2872.

Dunsmoor, J.E., Murty, V.P., Davachi, L., and Phelps, E.A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. Nature *520*, 345–348.

Dunsmoor, J.E., Kroes, M.C.W., Moscatelli, C.M., Evans, M.D., Davachi, L., and Phelps, E.A. (2018). Event segmentation protects emotional memories from competing experiences encoded close in time. Nat. Hum. Behav. *2*, 291–299.

Fenker, D.B., Frey, J.U., Schuetze, H., Heipertz, D., Heinze, H.J., and Duzel, E. (2008). Novel scenes improve recollection and recall of words. J. Cogn. Neurosci. *20*, 1250–1265.

Frey, U., and Morris, R.G.M. (1997). Synaptic tagging and long-term potentiation. Nature *385*, 533–536.

Frossard, J., and Renaud, O. (2019). permuco: Permutation Tests for Regression, Repeated Measures) ANOVA/ANCOVA and Comparison of Signals.

Gershman, S.J., Schapiro, A.C., Hupbach, A., and Norman, K.A. (2013). Neural context reinstatement predicts memory misattribution. J. Neurosci. *33*, 8590–8595.

Hennings, A.C., McClay, M., Lewis-Peacock, J.A., and Dunsmoor, J.E. (2020). Contextual reinstatement promotes extinction generalization in healthy adults but not PTSD. Neuropsychologia *147*, 107573.

Howard, M.W., and Kahana, M.J. (2002a). A distributed representation of temporal context. J. Math. Psychol. *46*, 269–299.

Howard, M.W., and Kahana, M.J. (2002b). When does semantic similarity help episodic retrieval? J. Mem. Lang. *46*, 85–98.

Johnson, M.K., Hashtroudi, S., and Lindsay, D.S. (1993). Source Monitoring. Psychol. Bull. *114*, 3–28.

Keller, N.E., and Dunsmoor, J.E. (2020). The effects of aversive-to-appetitive counterconditioning on implicit and explicit fear memory. Learn. Mem. *27*, 12–19.

Kensinger, E.A. (2009). Remembering the details: Effects of emotion. Emot. Rev. *1*, 99–113.

Kensinger, E.A., and Schacter, D.L. (2006). Reality monitoring and memory distortion: Effects of negative, arousing content. Mem. Cogn. *34*, 251–260.

LaBar, K.S., and Cabeza, R. (2006). Cognitive neuroscience of emotional memory. Nat. Rev. Neurosci. *7*, 54–64.

Lawrence, M.A. (2016). ez: Easy Analysis and Visualization of Factorial Experiments.

Lei, Y., Wang, J., Dou, H., Qiu, Y., and Li, H. (2019). Influence of typicality in category-based fear generalization: Diverging evidence from the P2 and N400 effect. Int. J. Psychophysiol. *135*, 12–20.

Mack, M.L., Love, B.C., and Preston, A.R. (2016). Dynamic updating of hippocampal object representations reflects new conceptual knowledge. Proc. Natl. Acad. Sci. U. S. A. *113*, 13203–13208.

Mather, M. (2007). Emotional Arousal and Memory Binding: An Object-Based Framework. Perspect. Psychol. Sci. *2*, 33–52.

Mather, M., Clewett, D., Sakaki, M., and Harley, C.W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. Behav. Brain Sci. *39*.

McGaugh, J.L. (2015). Consolidating memories. Annu. Rev. Psychol. *66*, 1–24.

Moncada, D., and Viola, H. (2007). Induction of long-term memory by exposure to novelty requires protein synthesis: Evidence for a behavioral tagging. J. Neurosci. *27*, 7476–7481.

Murty, V.P., Ritchey, M., Adcock, R.A., and LaBar, K.S. (2010). FMRI studies of successful emotional memory encoding: A quantitative meta-analysis. Neuropsychologia *48*, 3459–3469.

Patil, A., Murty, V.P., Dunsmoor, J.E., Phelps, E.A., and Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. Learn. Mem. *24*, 65–69.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2012). Scikit-learn: Machine Learning in Python. J.

Mach. Learn. Res. *12*, 2825–2830.

2 Pernet, C.R., Wilcox, R., and Rousselet, G.A. (2013). Robust correlation analyses: False positive and power validation using a new open source matlab toolbox. Front. Psychol. *3*, 606.

4 Quiroga, R.Q. (2012). Concept cells: the building blocks of declarative memory functions. Nat. Rev. Neurosci. *13*, 587–597.

6 Ramirez Butavand, D., Hirsch, I., Tomaiuolo, M., Moncada, D., Viola, H., and Ballarini, F. (2020). Novelty Improves the Formation and Persistence of Memory in a Naturalistic School
8 Scenario. Front. Psychol. *11*, 48.

Redondo, R.L., and Morris, R.G.M. (2011). Making memories last: The synaptic tagging and
10 capture hypothesis. Nat. Rev. Neurosci. *12*, 17–30.

Rimmele, U., Davachi, L., and Phelps, E.A. (2012). Memory for time and place contributes to
12 enhanced confidence in memories for emotional events. Emotion *12*, 834–846.

Schmidt, K., Patnaik, P., and Kensinger, E.A. (2011). Emotion's influence on memory for spatial
14 and temporal context. Cogn. Emot. *25*, 229–243.

Shapiro, S.S., and Wilk, M.B. (1965). An Analysis of Variance Test for Normality (Complete
16 Samples). Biometrika *52*, 591.

Sharot, T., and Phelps, E.A. (2004). How arousal modulates memory: Disentangling the effects
18 of attention and retention. Cogn. Affect. Behav. Neurosci. *4*, 294–306.

Sharot, T., and Yonelinas, A.P. (2008). Differential time-dependent effects of emotion on
20 recollective experience and memory for contextual information. Cognition *106*, 538–547.

Sharot, T., Delgado, M.R., and Phelps, E.A. (2004). How emotion enhances the feeling of
22 remembering. Nat. Neurosci. *7*, 1376–1380.

Smith, S.M., and Vela, E. (2001). Environmental context-dependent memory: A review and
24 meta-analysis. Psychon. Bull. Rev. *8*, 203–220.

Struyf, D., Hermans, D., and Vervliet, B. (2018). Maximizing the generalization of fear extinction:
26 Exposures to a peak generalization stimulus. Behav. Res. Ther. *111*, 1–8.

Takashima, A., van der Ven, F., Kroes, M.C.W., and Fernández, G. (2016). Retrieved emotional
28 context influences hippocampal involvement during recognition of neutral memories. Neuroimage *143*, 280–292.

30 Takeuchi, T., Duszkiewicz, A.J., Sonneborn, A., Spooner, P.A., Yamasaki, M., Watanabe, M., Smith, C.C., Fernández, G., Deisseroth, K., Greene, R.W., et al. (2016). Locus coeruleus and
32 dopaminergic consolidation of everyday memory. Nature *537*, 357–362.

Talmi, D., Lohnas, L.J., and Daw, N.D. (2019). A retrieved context model of the emotional
34 modulation of memory. Psychol. Rev. *126*, 455–485.

Tambini, A., Rimmele, U., Phelps, E.A., and Davachi, L. (2017). Emotional brain states carry
36 over and enhance future memory formation. Nat. Neurosci. *20*, 271–278.

Tulving, E. (2002). Episodic memory: From mind to brain.

38 Tulving, E., and Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychol. Rev. *80*, 352–373.

Vallat, R. (2018). Pingouin: statistics in Python. J. Open Source Softw. *3*, 1026.

2   Vargha, A., Delaney, H.D., and Vargha, A. (2000). A Critique and Improvement of the "CL" Common Language Effect Size Statistics of McGraw and Wong. J. Educ. Behav. Stat. *25*, 101.

4   Wang, B., and Fu, X. (2010). Gender differences in the effects of post-learning emotion on consolidation of item memory and source memory. Neurobiol. Learn. Mem. *93*, 572–580.

6   Wang, S.H., Redondo, R.L., and Morris, R.G.M. (2010). Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. Proc. Natl.
8   Acad. Sci. U. S. A. *107*, 19537–19542.

Yonelinas, A.P., and Ritchey, M. (2015). The slow forgetting of emotional episodic memories:
10   An emotional binding account. Trends Cogn. Sci. *19*, 259–267.

12

Supplementary Information for:

2

Emotional learning retroactively enhances item memory but distorts source attribution

4   Augustin C. Hennings, Jarrod A. Lewis-Peacock, Joseph E. Dunsmoor

Correspondence to: Joseph E. Dunsmoor

6   **Email**: joseph.dunsmoor@austin.utexas.edu

8   **This PDF file includes:**

Supplementary Tables 1-5

10   Supplementary Figure 1

12

	Temporal Context	CS Type	Mean	95% CI	Std. Error
Corrected Recognition	pre-conditioning	CS+	0.4106	[0.3648, 0.4579]	0.0239
		CS-	0.3560	[0.3079, 0.4051]	0.0251
	fear conditioning	CS+	0.4819	[0.4153, 0.5477]	0.0343
		CS-	0.3023	[0.2417, 0.3639]	0.0319
	post-conditioning	CS+	0.3023	[0.2426, 0.3644]	0.0311
		CS-	0.2181	[0.1769, 0.2616]	0.0224
False Alarm Rate	-	CS+	0.1403	[0.1037, 0.1810]	0.0199
	-	CS-	0.1218	[0.0861, 0.1611]	0.0194

2 **Supplementary Table 1. Mean corrected recognition memory and false alarm rate.**

4

Temporal Context	CS Type	DO	MO	MN	DN
Pre-conditioning	CS+	0.551	0.226	0.134	0.089
	CS-	0.478	0.257	0.172	0.093
Fear conditioning	CS+	0.622	0.227	0.094	0.056
	CS-	0.424	0.254	0.217	0.106
Post-extinction	CS+	0.443	0.268	0.181	0.109
	CS-	0.34	0.276	0.24	0.144
Novel lures	CS+	0.157	0.217	0.289	0.337
	CS-	0.134	0.199	0.307	0.36

6

Supplementary Table 2. Mean proportion of recognition memory responses. DO, definitely
8 old; MO, maybe old; MN, maybe new; DN, definitely new.

2

4

Temporal Context	CS Type	Response	Proportion	95% CI	P	
pre-conditioning	CS+	pre	0.2684	[0.2145, 0.3235]	0.0252	*
		cond.	0.4645	[0.3946, 0.5343]	0.0001	***
		post	0.2672	[0.2194, 0.3186]	0.0146	*
	CS-	pre	0.3284	[0.2819, 0.3775]	0.8544	
		cond.	0.3272	[0.2757, 0.3787]	0.8370	
		post	0.3444	[0.2978, 0.3934]	0.6678	
fear conditioning	CS+	pre	0.2108	[0.1605, 0.2647]	0.0001	***
		cond.	0.5821	[0.5086, 0.6556]	0.0001	***
		post	0.2071	[0.1569, 0.2659]	0.0004	***
	CS-	pre	0.3125	[0.2537, 0.3701]	0.5046	
		cond.	0.3309	[0.2745, 0.3897]	0.9252	
		post	0.3566	[0.3039, 0.4105]	0.3950	
post-conditioning	CS+	pre	0.2120	[0.1618, 0.2647]	0.0001	***
		cond.	0.4657	[0.4069, 0.5270]	0.0001	***
		post	0.3223	[0.2671, 0.3811]	0.7158	
	CS-	pre	0.2966	[0.2488, 0.3456]	0.1416	
		cond.	0.3137	[0.2770, 0.3505]	0.3032	
		post	0.3897	[0.3419, 0.4375]	0.0170	*

6 **Supplementary Table 3. Mean temporal context source memory data.**

Temporal Context	CS Type	Response	Beta	95% CI	P	
pre-conditioning	CS+	pre	0.0869	[-0.2138, 0.3861]	0.5656	
		cond.	0.4801	[0.2004, 0.7620]	0.0014	**
		post	-0.5671	[-0.8346, -0.3007]	0.0001	***
	CS-	pre	-0.1373	[-0.3929, 0.1263]	0.2936	
		cond.	0.3453	[0.1346, 0.5645]	0.0018	**
		post	-0.2080	[-0.4173, 0.0030]	0.0546	~
fear conditioning	CS+	pre	-0.1794	[-0.4909, 0.1294]	0.2634	
		cond.	0.6525	[0.4102, 0.9041]	0.0001	***
		post	-0.4731	[-0.7714, -0.1895]	0.0012	**
	CS-	pre	-0.2703	[-0.6053, 0.0628]	0.1180	
		cond.	0.6773	[0.3748, 0.9787]	0.0001	***
		post	-0.4070	[-0.7058, -0.1144]	0.0052	**
post-conditioning	CS+	pre	-0.1695	[-0.4961, 0.1404]	0.2990	
		cond.	0.5461	[0.234, 0.8563]	0.0004	***
		post	-0.3766	[-0.6921, -0.0679]	0.0154	*
	CS-	pre	-0.1893	[-0.4694, 0.0792]	0.1666	
		cond.	0.4660	[0.2659, 0.6884]	0.0001	***
		post	-0.2766	[-0.5447, -0.0186]	0.0354	*

2 **Supplementary Table 4. Temporal context memory logistic regression data.**

Temporal Context	CS Type	Mean	95% CI	P	
pre-conditioning	CS+	0.0796	[-0.0033, 0.1597]	0.0576	~
	CS-	0.1181	[0.0212, 0.2221]	0.0142	*
fear conditioning	CS+	0.1666	[0.0732, 0.2591]	0.0006	***
	CS-	0.0937	[0.0004, 0.197]	0.0488	*
post-conditioning	CS+	0.1741	[0.0733, 0.2764]	0.0002	***
	CS-	0.1293	[0.0182, 0.2622]	0.0202	*

2   **Supplementary Table 5. Typicality logistic regression data.**

Mean difference in typicality for each stimulus
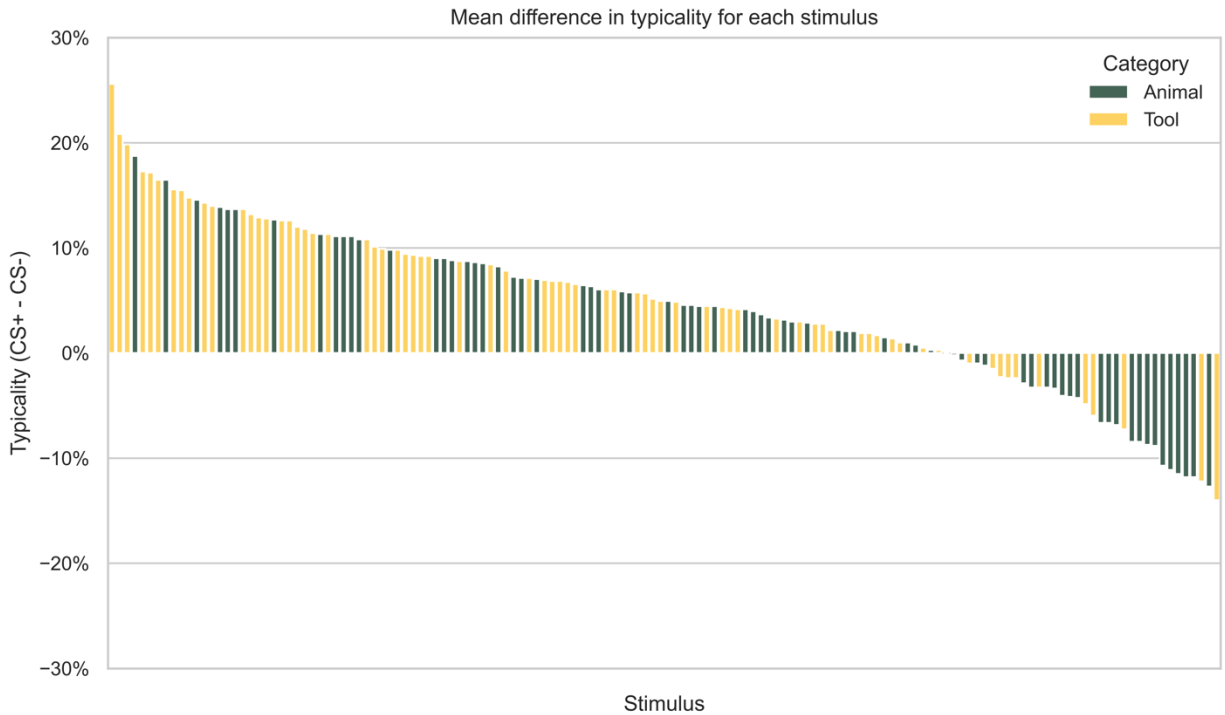
2 **Supplementary Figure 1.** Aversive learning enhances perceived typicality for stimuli. Mean difference in typicality for each stimulus (CS+ - CS-). This difference constitutes a between
4 groups analysis in which we subtracted the mean typicality response for each exemplar (e.g., a rabbit) when it was used as a CS+ versus a CS-, and is intended for illustration purposes only.
6 All 144 stimuli used during memory encoding are shown (unlabeled) and arranged by the difference in typicality between when the same exemplar was used as a CS+ or a CS-. The
8 particular encoding phase (pre-conditioning, fear conditioning, post-conditioning) each exemplar was presented was counterbalanced across subjects. Values reflect % difference of the 7-point
10 scale (14% = 1 point difference).

12