Emotional learning selectively and retroactively strengthens memories for related events

Joseph E. Dunsmoor[1], Vishnu P. Murty[1], Lila Davachi[1] & Elizabeth A. Phelps[1,2]

Neurobiological models of long-term memory propose a mechanism by which initially weak memories are strengthened through subsequent activation that engages common neural pathways minutes to hours later[1]. This synaptic tag-and-capture model has been hypothesized to explain how inconsequential information is selectively consolidated following salient experiences. Behavioural evidence for tag-and-capture is provided by rodent studies in which weak early memories are strengthened by future behavioural training[2,3]. Whether a process of behavioural tagging occurs in humans to transform weak episodic memories into stable long-term memories is unknown. Here we show, in humans, that information is selectively consolidated if conceptually related information, putatively represented in a common neural substrate, is made salient through an emotional learning experience. Memory for neutral objects was selectively enhanced if other objects from the same category were paired with shock. Retroactive enhancements as a result of emotional learning were observed following a period of consolidation, but were not observed in an immediate memory test or for items strongly encoded before fear conditioning. These findings provide new evidence for a generalized retroactive memory enhancement, whereby inconsequential information can be retroactively credited as relevant, and therefore selectively remembered, if conceptually related information acquires salience in the future.

People are motivated to remember the episodic details of emotional events, because this information is useful for predicting and controlling important events in the future[4,5]. In contrast, there is often little motivation to remember insignificant details we accumulate throughout the day, since much of this information is not associated with anything particularly meaningful. We do not always know, however, when a meaningful event will occur. From an adaptive memory perspective it is therefore critical that seemingly inconsequential details be stored in memory, at least temporarily, in the event that this information acquires relevance some time later. In this way, initially weak memories can be strengthened if this information later gains meaning. However, since we rarely encounter the same exact stimuli in the same exact situations it is advantageous for memories of other closely related information, encoded before a meaningful event, to be remembered as well. Such a mechanism could explain how a highly emotional event enhances memory for a host of details encoded earlier that, at the time, did not appear to hold any significance. Here, we provide evidence of a generalized retroactive memory enhancement in humans that is selective to information conceptually related to a future emotional event.

For episodic details to persist in long-term memory requires memory stabilization through the process of consolidation. A neurobiological account of memory consolidation has proposed a synaptic tag-and-capture mechanism whereby new memories that are initially weak and unstable are tagged for later stabilization by long-term potentiation (LTP) processes[1]. This mechanism has been extended to the domain of hippocampus-dependent learning in rats to explain how weak behavioural training that would otherwise be forgotten will endure in memory following a new behavioural experience (for example, exposure to novelty)—an effect referred to as behavioural tagging[2,3,6,7].

Whether behavioural tagging occurs in human episodic memory is unknown. Evidence for such an effect would require that memory for older events that are related to subsequent experiences is selectively enhanced while other unrelated information encoded at the same time should not receive a retroactive memory benefit. While prior studies have shown post-encoding modulation of memory consolidation with increases in stress and arousal[8,9], these demonstrations do not provide evidence of specificity. Another strong test of this hypothesized process is to mitigate the potential for selective rehearsal by presenting information in the absence of any motivation or instruction to remember (incidental encoding) and conducting a surprise memory test. Finally, models of behavioural tagging predict memory strengthening for weak encoding, but not strong encoding[6,7,10,11]. Thus, a task designed to retroactively boost relatively weak episodic memories should not retroactively benefit memories that were already strongly encoded.

Taking these criteria into consideration, we investigated whether information is selectively remembered if conceptually related information is later made salient through an amygdala-dependent learning task; that is, a trial-unique form of Pavlovian fear conditioning[12,13]. The encoding session occurred in three phases on the same day (Fig. 1). In phase 1, subjects classified 60 distinct basic-level objects as animals or tools (30 each). Shock electrodes were not attached during phase 1 and there was no explicit motivation or instruction to remember any of the pictures. Shortly thereafter, in phase 2, electric shock electrodes were attached and 30 novel images from one category (conditioned stimulus or CS$^+$, animals or tools, counterbalanced) were paired with a shock (unconditioned stimulus) to the right wrist at a reinforcement rate of 66%, while 30 novel images from the other category (CS$^-$, tools or animals, respectively) were unpaired. Skin conductance responses were acquired during fear conditioning to evaluate discriminatory fear learning. After conditioning, in phase 3, electric shock electrodes were removed and subjects classified additional images of 30 animals and 30 tools. Surprise recognition memory tests were then administered after either a 24-h delay, a 6-h delay, or immediately after phase 3 (see Methods for additional experimental details). The use of separate object categories provides the ability to test for selective consolidation in a within-subjects design. That is, we can assess whether fear conditioning preferentially enhances long-term memory for items related to the CS$^+$ but encoded before the conditioning experience, before any knowledge that related information would acquire future salience.

Significant physiological evidence of fear conditioning in phase 2, as assessed with greater skin conductance responses to the CS$^+$ versus the CS$^-$ category exemplars, was observed in all groups (Extended Data Fig. 1 and Methods). Recognition memory was calculated using corrected recognition (number of hits minus the number of false alarms to the corresponding category). An ANOVA with CS (CS$^+$, CS$^-$) and phase (pre-conditioning, conditioning, post-conditioning) as repeated measures, and retrieval group (24 h, 6 h, immediate) as between-subjects factor, revealed a main effect of CS ($F_{1,86} = 18.82$, $P < 0.001$, $\eta_p^2 = 0.18$), phase ($F_{2,85} = 29.35$, $P < 0.001$, $\eta_p^2 = 0.36$), and group ($F_{2,86} = 11.82$, $P < 0.001$, $\eta_p^2 = 0.22$), as well as a significant phase × group ($F_{4,172} = 4.49$,

[1]Department of Psychology and Center for Neural Sciences, New York University, New York, New York 10003, USA. [2]Nathan Kline Institute, Orangeburg, New York 10962, USA.
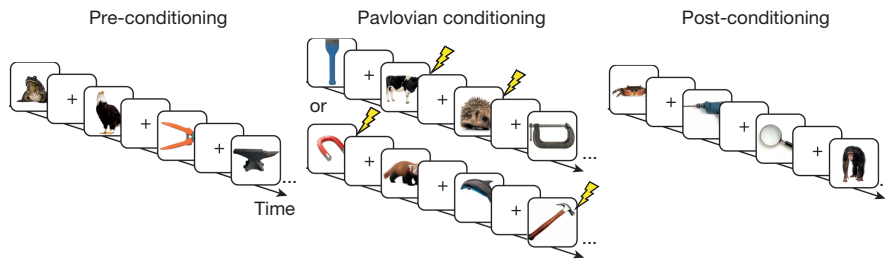
Figure 1 | Incidental encoding paradigm and example stimuli. Adult human subjects viewed 90 basic-level exemplars of animals and tools before, during and after fear conditioning. Before and after fear conditioning, subjects classified each object as an animal or a tool. During conditioning, electric shocks were paired with 20 out of 30 animal or tool pictures (counterbalanced between subjects) while subjects rated shock expectancy. A surprise recognition memory test was administered 24 h ($n = 30$), 6 h ($n = 30$), or immediately ($n = 29$) after encoding. Lightning bolts denote electric shocks.

$P = 0.002$, $\eta_p^2 = 0.09$) interaction. Follow-up planned ANOVAs and t-tests were conducted separately for the three retrieval groups.

The 24-h retrieval group (Fig. 2a and Extended Data Fig. 2) showed a main effect of CS ($F_{1,29} = 18.76$, $P < 0.001$, $\eta_p^2 = 0.39$) and phase ($F_{2,28} = 9.35$, $P = 0.001$, $\eta_p^2 = 0.40$). Follow up t-tests revealed that recognition memory was enhanced for CS$^+$ items encoded during fear conditioning ($t_{29} = 3.47$, $P = 0.002$, $d_{av} = 0.53$ (see Methods for an explanation of d_{av})), replicating previous findings[9]. This memory benefit extended to CS$^+$ exemplars encoded after fear conditioning, when the shock electrodes were unattached ($t_{29} = 3.42$, $P = 0.002$, $d_{av} = 0.58$), suggesting that selective effects of conditioning on subsequent memory can operate prospectively. Critically, a retroactive memory enhancement for CS$^+$ items was also observed. Memory was significantly stronger for items conceptually related to the CS$^+$ versus items related to the CS$^-$ encoded before conditioning ($t_{29} = 2.48$, $P = 0.019$, $d_{av} = 0.41$), suggesting that weak memories from the pre-conditioning session were bolstered once conceptually related information acquired emotional relevance. There were no differences in false alarms between CS conditions ($P = 0.57$).

At 6-h retrieval (Fig. 2b), there was a main effect of CS ($F_{1,29} = 6.93$, $P = 0.01$, $\eta_p^2 = 0.19$) but no effect of phase ($P = 0.11$). The CS × phase interaction was significant ($F_{2,28} = 3.46$, $P = 0.05$, $\eta_p^2 = 0.19$). Follow-up t-tests showed significantly greater memory for CS$^+$ versus CS$^-$ items encoded during pre-conditioning ($t_{29} = 2.41$, $P = 0.02$, $d_{av} = 0.40$) and fear conditioning ($t_{29} = 2.80$, $P = 0.009$, $d_{av} = 0.48$), replicating results obtained from 24-h retrieval. No differences between CS$^+$ and CS$^-$ memory emerged at post-conditioning ($P = 0.52$), and there were no differences in false alarms between CS conditions ($P = 0.95$). This result indicates that fear-conditioning-mediated retroactive memory enhancements emerge by 6 h and are not dependent on sleep consolidation.

At immediate retrieval (Fig. 2c), there was no main effect of CS ($P = 0.17$), but there was a main effect of phase ($F_{2,27} = 20.32$, $P < 0.001$, $\eta_p^2 = 0.60$). Follow-up t-tests showed significantly greater memory for CS$^+$ versus CS$^-$ items encoded during fear conditioning ($t_{28} = 2.14$, $P = 0.04$, $d_{av} = 0.50$). However, there was no difference in recognition memory between CS$^+$ and CS$^-$ items encoded during pre-conditioning ($P = 0.97$) or post-conditioning ($P = 0.21$), and no differences in false alarms between CS conditions ($P = 0.74$). Importantly, this result suggests that fear-conditioning-mediated retroactive memory enhancement requires a period of consolidation. In order to directly assess whether the retroactive memory enhancement differed for delayed versus immediate retrieval, a memory difference score (corrected recognition for CS$^+$ items minus CS$^-$ items) was calculated for all groups (Fig. 3). A comparison of delayed (24 and 6 h) versus immediate retrieval revealed significantly greater memory for CS$^+$ versus CS$^-$ items encoded during pre-conditioning in the delayed groups relative to the immediate retrieval group ($t_{87} = 1.77$, $P = 0.04$, one-tailed, $d = 0.38$). In contrast to results from pre-conditioning, a comparison of post-conditioning memory between same-day (immediate and 6 h) versus next-day (24 h) retrieval revealed significantly greater memory for CS$^+$ versus CS$^-$ items encoded during post-conditioning in the next-day group relative to the same-day groups ($t_{87} = 1.66$, $P = 0.05$, one-tailed, $d = 0.38$).

Models of behavioural tagging predict retroactive effects on weakly encoded memories, but no effect for strongly encoded memories[6,7,10,11]. To test whether strong encoding presents a boundary condition for retroactive enhancements of episodic memory, a separate group was
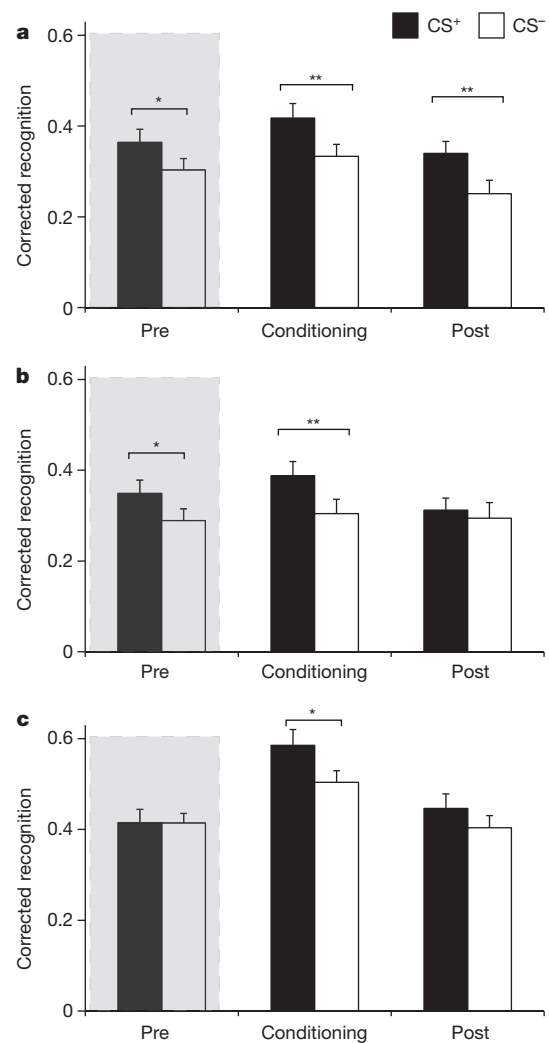
Figure 2 | Recognition memory performance. Memory at 24-h (**a**), 6-h (**b**), and immediate (**c**) retrieval showed enhanced corrected recognition memory for items from the CS$^+$ versus the CS$^-$ category encoded during fear conditioning in all groups. However, memory was only retroactively enhanced for CS$^+$ items encoded during pre-conditioning following a 24-h or a 6-h delay. The shaded area highlights retroactive memory for items that preceded fear conditioning. CS$^+$, conditioned stimuli from the object category with exemplars paired with shock; CS$^-$, conditioned stimuli from the object category with exemplars never paired with shock. Error bars are s.e.m. *$P < 0.05$, **$P < 0.01$, two-tailed t-tests.
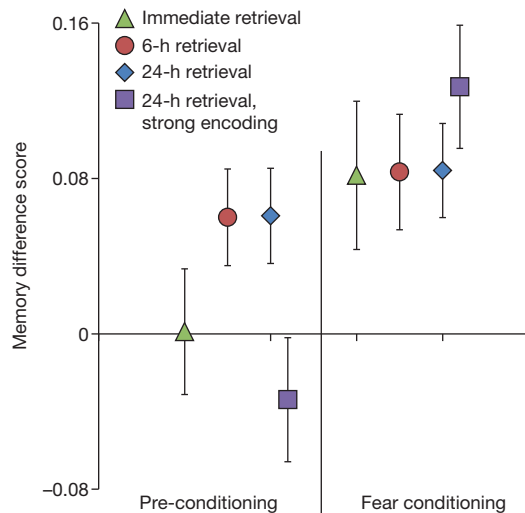
Figure 3 | Recognition memory difference scores. Corrected recognition difference scores (CS$^+$ minus CS$^-$) highlight that selective retroactive memory enhancements emerged at delay, but not immediate test, and not in subjects for whom pre-conditioning memory was strong before fear conditioning. Memory enhancements during fear conditioning were observed in all four experimental groups. Error bars are s.e.m.

shown each stimulus three times during pre-conditioning before fear conditioning. A surprise memory test was conducted 24 h later. An ANOVA showed a trend for CS ($P = 0.073$), an effect of phase ($F_{1,29} = 27.07$, $P < 0.001$, $\eta_p^2 = 0.49$), and a significant CS $\times$ phase interaction ($F_{1,29} = 17.04$, $P < 0.001$, $\eta_p^2 = 0.37$). Follow-up t-tests showed that this interaction was driven by greater corrected recognition memory for CS$^+$ (0.57 $\pm$ 0.03 (mean $\pm$ standard error)) than CS$^-$ (0.45 $\pm$ 0.03) during fear conditioning ($t_{29} = 4.02$, $P = 0.02$, $d_{av} = 0.73$), and no difference between CS$^+$ (0.66 $\pm$ 0.04) and CS$^-$ (0.69 $\pm$ 0.03) items encoded during pre-conditioning ($P = 0.29$) (Extended Data Fig. 3). A direct comparison between the 24-h weak- and 24-h strong-encoding groups revealed, as predicted, significantly greater overall (CS$^+$ and CS$^-$) memory for the strong-encoding (0.68 $\pm$ 0.03) versus weak-encoding (0.33 $\pm$ 0.02) group during pre-conditioning ($t_{58} = 3.87$, $P < 0.001$, $d = 0.99$), but a comparison of the memory difference score (CS$^+$ minus CS$^-$; Fig. 3) demonstrated a selective memory enhancement for CS$^+$ items in the 24-h weak-encoding group only ($t_{58} = 2.35$, $P = 0.01$, $d = 0.61$) (see Methods for additional analyses).

We found that memories for neutral information can be enhanced by a future emotional event that involves conceptually related material. The use of two category domains with relatively well-delineated neural substrates[14] allows us to speculate on a potential neurobiological mechanism mediating these effects. A recent neuroimaging investigation[13] showed that fear conditioning at the categorical level with animals and tools (akin to phase 2 from these experiments) modulates activity in category-selective regions in the extrastriate visual cortex; that is, activity in category-selective regions is enhanced in subjects for whom novel pictures of animals (or tools) predict shock. In the context of the present study, encoding during the pre-conditioning classification task may have set a weak learning tag in the hippocampus and these category-selective regions in the occipitotemporal cortex. Fear-conditioning-induced modulation of category-selective cortex and the hippocampus, via the amygdala or other regions involved in emotional learning circuitry, may then enhance related memories and possibly selectively prune unrelated memories[15]. Although these results are consistent with a putative tag-and-capture mechanism, whether such a mechanism explains the behavioural effect shown here requires future research. A consolidation mechanism is supported by the observation that memory enhancements for pre-conditioning were not seen in an immediate memory test. Notably, retroactive enhancements were evident after 6 h, in line

with studies showing that arousal-mediated consolidation effects are dependent on time, but not dependent on sleep[16,17]. This is in contrast to research showing selective retention for items retroactively made relevant through explicit instructions to remember, which finds effects only after a period of sleep consolidation[18,19].

This generalized retroactive memory enhancement can also be distinguished from prior studies of global post-encoding increases in consolidation through administration of stress or arousal[8,9], as emotional learning selectively enhanced memory for neutral items associated with that category, but not other neutral content encoded at the same time. By virtue of presenting information before Pavlovian conditioning, we can also disentangle enhanced attention at the time of encoding induced by the anticipation of shock from post-encoding consolidation processes. That is, during phase 1 there was no chance of receiving shocks (the shock leads were not attached), and no details had been provided to the subject about the contingencies of shock administration for later phases of the experiment (see Methods for further details). These results are also different from generalization that involves overlapping representations of cues pre-associated before reinforcement; for example, acquired equivalence[20]. In the present study, information presented at each phase of encoding is related at the conceptual level, but is never repeated or directly combined with information presented at another phase of incidental encoding.

Notably, while a retroactive memory benefit was shown after 24-h and 6-h delays, a proactive memory benefit was only demonstrated after 24 h. This finding was unexpected, and indicates that retroactive and proactive arousal-mediated memory enhancements are separable and perhaps rely on different mechanisms. In a potentially analogous finding[21], the ability to make inferential judgments regarding previously learned relational knowledge was reported to increase following a delay, and be further boosted following sleep. Whether the proactive memory enhancement in this study relies on a period of sleep consolidation is an intriguing possibility that may help dissociate mechanisms supporting retroactive versus proactive emotional memory effects.

In conclusion, our work provides new evidence for selective consolidation of information conceptually related to a future meaningful event. These findings support an implication proposed previously[1] in the formulation of the synaptic tag-and-capture mechanism, that late-phase LTP of synaptic activity could explain enhanced memories for seemingly insignificant details surrounding emotional events. An intriguing implication of this finding concerns the adaptive nature of episodic memory. Specifically, humans and other animals continuously monitor the environment, accumulating countless details. Much of this information is forgotten. However, meaningful events can selectively preserve memory for previously encountered information that seemed insignificant at the time it was encoded. Whether such a mechanism contributes to persistent intrusive memories and overgeneralization of fear characteristic of trauma and stress-related disorders merits further empirical research.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Frey, U. & Morris, R. G. Synaptic tagging and long-term potentiation. *Nature* **385**, 533–536 (1997).
2. Ballarini, F., Moncada, D., Martinez, M. C., Alen, N. & Viola, H. Behavioral tagging is a general mechanism of long-term memory formation. *Proc. Natl Acad. Sci. USA* **106**, 14599–14604 (2009).
3. de Carvalho Myskiw, J., Benetti, F. & Izquierdo, I. Behavioral tagging of extinction learning. *Proc. Natl Acad. Sci. USA* **110**, 1071–1076 (2013).
4. LaBar, K. S. & Cabeza, R. Cognitive neuroscience of emotional memory. *Nature Rev. Neurosci.* **7**, 54–64 (2006).
5. Lisman, J., Grace, A. A. & Duzel, E. A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends Neurosci.* **34**, 536–547 (2011).
6. Wang, S. H., Redondo, R. L. & Morris, R. G. Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. *Proc. Natl Acad. Sci. USA* **107**, 19537–19542 (2010).

7. Moncada, D. & Viola, H. Induction of long-term memory by exposure to novelty requires protein synthesis: evidence for a behavioral tagging. *J. Neurosci.* **27,** 7476–7481 (2007).
8. McGaugh, J. L. The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.* **27,** 1–28 (2004).
9. Cahill, L. & McGaugh, J. L. Mechanisms of emotional arousal and lasting declarative memory. *Trends Neurosci.* **21,** 294–299 (1998).
10. Moncada, D., Ballarini, F., Martinez, M. C., Frey, J. U. & Viola, H. Identification of transmitter systems and learning tag molecules involved in behavioral tagging during memory formation. *Proc. Natl Acad. Sci. USA* **108,** 12931–12936 (2011).
11. da Silva, B. M., Bast, T. & Morris, R. G. Spatial memory: behavioral determinants of persistence in the watermaze delayed matching-to-place task. *Learn. Mem.* **21,** 28–36 (2014).
12. Dunsmoor, J. E., Martin, A. & LaBar, K. S. Role of conceptual knowledge in learning and retention of conditioned fear. *Biol. Psychol.* **89,** 300–305 (2012).
13. Dunsmoor, J. E., Kragel, P. A., Martin, A. & LaBar, K. S. Aversive learning modulates cortical representations of object categories. *Cereb. Cortex* **24,** 2859–2872 (2014).
14. Martin, A. The representation of object concepts in the brain. *Annu. Rev. Psychol.* **58,** 25–45 (2007).
15. Stickgold, R. & Walker, M. P. Sleep-dependent memory triage: evolving generalization through selective processing. *Nature Neurosci.* **16,** 139–145 (2013).
16. Park, J. Effect of arousal and retention delay on memory: a meta-analysis. *Psychol. Rep.* **97,** 339–355 (2005).
17. Kleinsmith, L. J. & Kaplan, S. Paired-associate learning as a function of arousal and interpolated interval. *J. Exp. Psychol.* **65,** 190–193 (1963).
18. Wilhelm, I. *et al.* Sleep selectively enhances memory expected to be of future relevance. *J. Neurosci.* **31,** 1563–1569 (2011).
19. van Dongen, E. V., Thielen, J. W., Takashima, A., Barth, M. & Fernandez, G. Sleep supports selective retention of associative memories based on relevance for future utilization. *PLoS ONE* **7,** e43426 (2012).
20. Shohamy, D. & Wagner, A. D. Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* **60,** 378–389 (2008).
21. Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D. & Walker, M. P. Human relational memory requires time and sleep. *Proc. Natl Acad. Sci. USA* **104,** 7723–7728 (2007).

Author Contributions J.E.D. designed and conducted the study. J.E.D. and V.P.M. analysed the data. J.E.D., V.P.M., L.D. and E.A.P. interpreted the results and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.D. (lila.davachi@nyu.edu) or E.A.P. (liz.phelps@nyu.edu).

METHODS

Participants. A total of 138 subjects were recruited to participate. Nineteen subjects were removed from the analysis for failure to return for the memory test ($n = 6$), failure to understand or follow the task instructions ($n = 7$), equipment failures with stimulus presentation software ($n = 3$), a failure to show any evidence of recognition memory above chance ($n = 2$), or indicating that the memory test was not a surprise ($n = 1$). The final sample included 119 subjects (Age = 23.42 ± 3.15 years (mean $\pm$ s.d.), 62 females). Subjects were assigned to 1 of 4 groups, immediate retrieval ($n = 29$, 16 females), 6-h retrieval ($N = 30$, 15 females), 24-h retrieval ($n = 30$, 20 females), or 24-h retrieval strong pre-conditioning encoding ($n = 30$, 11 females). Subjects in the immediate and 24-h retrieval groups were randomly assigned. The 6-h and 24-h strong-encoding groups were run as follow-up studies, and group assignment was not determined by randomization. Sample size was based on prior studies of categorical fear learning[12,13]. No statistical method was used to predetermine sample size. All subjects provided written informed consent approved by the University Committee on Activities Involving Human Subjects at New York University.

Behavioural paradigm and stimulus materials for 24-h, 6-h and immediate retrieval groups. The study involved two experimental sessions: incidental encoding and a surprise recognition memory test. The incidental-encoding session included 3 phases: pre-conditioning, fear conditioning, and post-conditioning. Each phase included 30 colour photographs of animals and 30 colour photographs of tools presented on a white background. Pictures were obtained from the website http://www.lifeonwhite.com or from publicly available resources on the internet. Each picture was a different basic-level exemplar with a different name; for example, there were not two different pictures of a dog. Stimulus order was counterbalanced across subjects and pseudo-randomized such that no more than 3 pictures from the same category appeared in a row.

During pre-conditioning, pictures were presented for 2.5 s with a 6 ± 2 s variable inter-trial interval that included a fixation cross on a blank background. The total duration of pre-conditioning was ~ 8.5 min. During pre-conditioning subjects made two-alternative forced-choice picture identifications ('animal' or 'tool'). Specifically, subjects were asked to classify each picture as either an animal or a tool by pressing the 1 or 2 button on a keypad on every trial. The buttons corresponding to animal and tool were counterbalanced across subjects.

Fear conditioning followed pre-conditioning ~ 5 min later. Between pre-conditioning and fear conditioning, shock leads were attached to the right wrist, and intensity was calibrated to a level deemed highly unpleasant, but not painful, using an ascending staircase procedure. Skin conductance response (SCR) leads were attached to the left palm. During fear conditioning, pictures were presented for 4.5 s with a variable inter-trial interval of 8 ± 2 s, which allowed time to measure SCRs before shocks occurred on CS^+ trials, and for SCRs to return to baseline after CS presentation. Shocks occurred on 20 out of 30 CS^+ trials at the end of the trial, co-terminating with the picture. The CS^+ trials paired with shock were counterbalanced between subjects. The total duration of fear conditioning was ~ 12 min. During fear conditioning, subjects made a two-alternative forced-choice shock expectancy rating (1 = shock, 2 = no shock). Specifically, subjects were asked to rate whether they expected the shock or not on every trial. Subjects were not instructed about the conditioned–unconditioned stimulus contingencies, and had to learn the category level association between the pictures and the shock through experience. Subjects were told explicitly that the button presses did not have any effect on whether or not the shock would occur, thus eliminating the chance for subjects to mistakenly attribute the outcome to their actions. The object categories serving as CS^+/CS^- were counterbalanced between subjects. After fear conditioning, the shock leads were removed and subjects were asked to rate the intensity of the shock on a scale from 1 (not at all unpleasant) to 10 (extremely unpleasant). The average rating was 6.17 (s.d. = 1.46), and there were no differences in mean intensity ratings between groups.

After fear conditioning the shock electrodes were removed. Post-conditioning occurred approximately 3 min after the end of fear conditioning. Procedures and instructions for post-conditioning were identical to those of pre-conditioning.

Recognition memory test procedures. The recognition memory test included the 90 CS^+ and 90 CS^- pictures seen the previous day, along with 90 new pictures of animals and 90 new pictures of tools (total of 360 pictures shown during the recognition memory test). The test was self-paced. Subjects rated whether each picture was new or old and their confidence by making 1 of 4 possible responses: definitely new, maybe new, maybe old, or definitely old. Memory responses were collapsed across confidence. Analysis focusing on high-confidence responses yielded similar results (all data are presented in Extended Data Tables 1–4). We performed our analysis on corrected recognition scores (hits minus false alarms) to account for differences in response criteria across participants. Data were normally distributed and variance was similar between groups. There were no differences in false alarms between CS categories (reported in main text).

Behavioural paradigm for strong encoding, 24-h retrieval. A separate group underwent a modified version of phase 1 encoding in which each stimulus ($n = 30$ animals, $n = 30$ tools) was presented 3 times each to strengthen memory for items encoded before fear conditioning. Trial order was randomized with the following constraints. First, no more than three images from the same object category appeared in a row. Second, each exemplar was presented twice during the first 120 trials, and once in the final 60 trials. Each picture was presented once during the final 60 trials to ensure that the lag between final stimulus presentation and conditioning was matched with the other protocols. To help ensure that the total duration of the experimental session was equivalent to the other groups, the inter-trial interval during phase 1 was reduced to 3.5 ± 0.5 s. Phase 1 was followed by the fear-conditioning protocol employed in the other groups (30 novel CS^+ and 30 novel CS^- trials, with 20/30 CS^+ trials paired with shock). As we were specifically interested in the effects on retroactive memory enhancements, we did not conduct a post-conditioning encoding session. This also helped keep the total time of the encoding session equivalent to the other experimental groups. The retrieval test for the strong-encoding group included the 60 CS^+ and 60 CS^- pictures seen the previous day, along with an equal number of new pictures from the CS^+ and CS^- categories (60 each).

Subject instructions and explicit knowledge regarding fear conditioning and memory test. Subjects were informed in advance that the study would involve electrical stimulation, and during informed consent each subject was told where the shock electrodes would be placed, and how the experimenter would calibrate the shock to a level they deemed highly unpleasant, but not painful. Importantly, no specific information was provided regarding the fear-conditioning phase before pre-conditioning, and shock leads were not attached during pre-conditioning. Consequently, even if subjects anticipated receiving shocks at a later phase of the experiment, this could not have a selective effect for one category of objects, since no details were provided regarding fear conditioning by this point of the task.

To assess whether subjects expected the surprise memory test, subjects in the 6-h retrieval group and the 24-h strong-encoding group were asked two questions when they returned for the memory test. First, they were asked, "Do you have any expectations of what this next task in the experiment will be: yes or no?" Subjects were then told that we would be conducting a test of their memory for the pictures they saw earlier, and were asked to indicate on a 5-point scale how surprised they were by a memory test, from 1 (I did not expect a memory test at all) to 5 (Yes, I knew there would be a memory test). The mean response was 2.63 (s.d. = 1.08). Only one subject responded "yes" to the first question and guessed correctly about a memory test. This was also the only subject to respond "5" on the second question. This subject was not included in the analysis.

Estimates of effect size. Effect sizes reported for ANOVAs in the manuscript are partial eta squared. For paired t-tests, we calculated Cohen's d using the mean difference score as the numerator and the average standard deviation of both repeated measures as the denominator, as suggested in ref. 22. This effect size is referred to in the text as d_{av}, where 'av' refers to the use of the average standard deviations in the calculation.

Shock and psychophysiology. A 200-ms shock was delivered to the right wrist using pre-gelled snap electrodes (BIOPAC EL508) connected to a Grass Medical Instruments stimulator (West Warwick, Rhode Island). SCR electrodes were placed on the hypothenar eminence of the palmar surface of the left hand using pre-gelled snap electrodes (BIOPAC EL509). Data were collected using a BIOPAC MP-100 System (Goleta, CA), and responses calculated using established criteria[23,24]. In brief, an SCR was considered related to CS presentation if the trough-to-peak deflection occurred 0.5–4.5 s following CS onset, lasted between 0.5 and 5.0 s, and was greater than 0.02 microsiemens (μS). Responses that did not fit these criteria were scored as zero. SCR values were obtained using a custom Matlab (The MathWorks, Inc.) script that extracted SCRs for each trial using the above criteria[25].

SCR results. SCRs were collected as a manipulation check that the fear-conditioning procedure effectively generated higher autonomic arousal on CS^+ trials than CS^- trials. SCR data was not analysed for 13 subjects due to equipment malfunction with the BIOPAC during data acquisition (24-h, $n = 4$; 6-h, $n = 3$; immediate, $n = 5$; 24-h strong, $n = 1$) and for 6 subjects due to an overall lack of measurable electrodermal responses (24-h, $n = 1$; 6-h, $n = 1$; immediate, $n = 2$; 24-h strong, $n = 2$). Paired t-tests showed enhanced SCRs to the CS^+ versus the CS^- in all four groups, and all P values were <0.0002, providing confirmation that the fear-conditioning manipulation was effective.

Supplementary memory analyses. To evaluate whether the memory enhancement observed for CS^+ versus CS^- items was different between the pre-conditioning and fear-conditioning phases, we compared the memory difference scores (CS^+ minus CS^-) between these two phases. This analysis was restricted to the two groups showing a selective CS^+ retroactive memory enhancement, the 24-h and 6-h delay groups. The memory difference score between pre-conditioning and fear conditioning was not different for either the 24-h ($P = 0.49$) or the 6-h ($P = 0.52$) delay

group. This analysis confirms that the CS^+ retroactive memory enhancement was not significantly different from the CS^+ fear-conditioning memory enhancement.

To ensure that the object categories serving as CS^+ and CS^- did not interact with memory effects, the object category subgroup (that is, animal CS^+/tool CS^-; tool CS^+/animal CS^-) was included as a covariate in a supplementary ANOVA. Subgroup did not interact with CS and phase for any group (all P values > 0.32).

As rodent studies of behavioural tagging show that the time interval between weak encoding and exposure to novelty can influence memory strength[3], we explored whether memory for items encoded during pre-conditioning (phase 1) were affected by the time relative to the start of fear conditioning (phase 2) in the 24-h retrieval group. For this analysis, items from pre-conditioning were binned according to tertiles corresponding to CS^+ (and CS^-) trials 0–10, 11–20 and 21–30. Tertiles roughly correspond to ∼14 to 11, ∼11 to 8, and ∼8 to 5 min before the start of fear conditioning, respectively. An ANOVA on the CS^+ minus CS^- memory difference score using tertiles as a factor revealed a significant linear effect ($F_{1,29} = 4.40$,

$P = 0.044$, $\eta_p^2 = 0.13$), such that memory difference between CS^+ and CS^- trials diminished from the first tertile (11 ± 0.03 (mean ± s.e.m.)), to the second tertile (0.08 ± 0.04), to the third tertile (0.03 ± 0.03). This result suggests that the time between weak episodic encoding and emotional learning may influence the strength of retroactive memory enhancements.

22. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Front. Psychol.* **4,** 863 (2013).
23. Schiller, D. *et al.* Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* **463,** 49–53 (2010).
24. Dunsmoor, J. E., Mitroff, S. R. & LaBar, K. S. Generalization of conditioned fear along a dimension of increasing fear intensity. *Learn. Mem.* **16,** 460–469 (2009).
25. Green, S. R., Kragel, P. A., Fecteau, M. E. & LaBar, K. S. Development and validation of an unsupervised scoring system (Autonomate) for skin conductance response analysis. *Int. J. Psychophysiol.* **91,** 186–193 (2014).
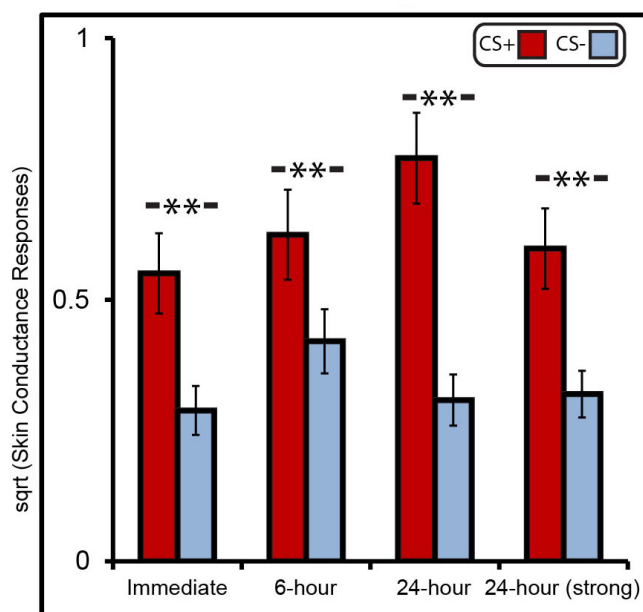
Extended Data Figure 1 | Skin conductance responses. Mean square-root-normalized skin conductive responses for the CS^+ and CS^- for each group. Results provide confirmation that the fear-conditioning manipulation was effective at generating higher arousal in CS^+ than CS^- trials. Error bars are s.e.m. $**P < 0.01$, two-tailed t-tests.
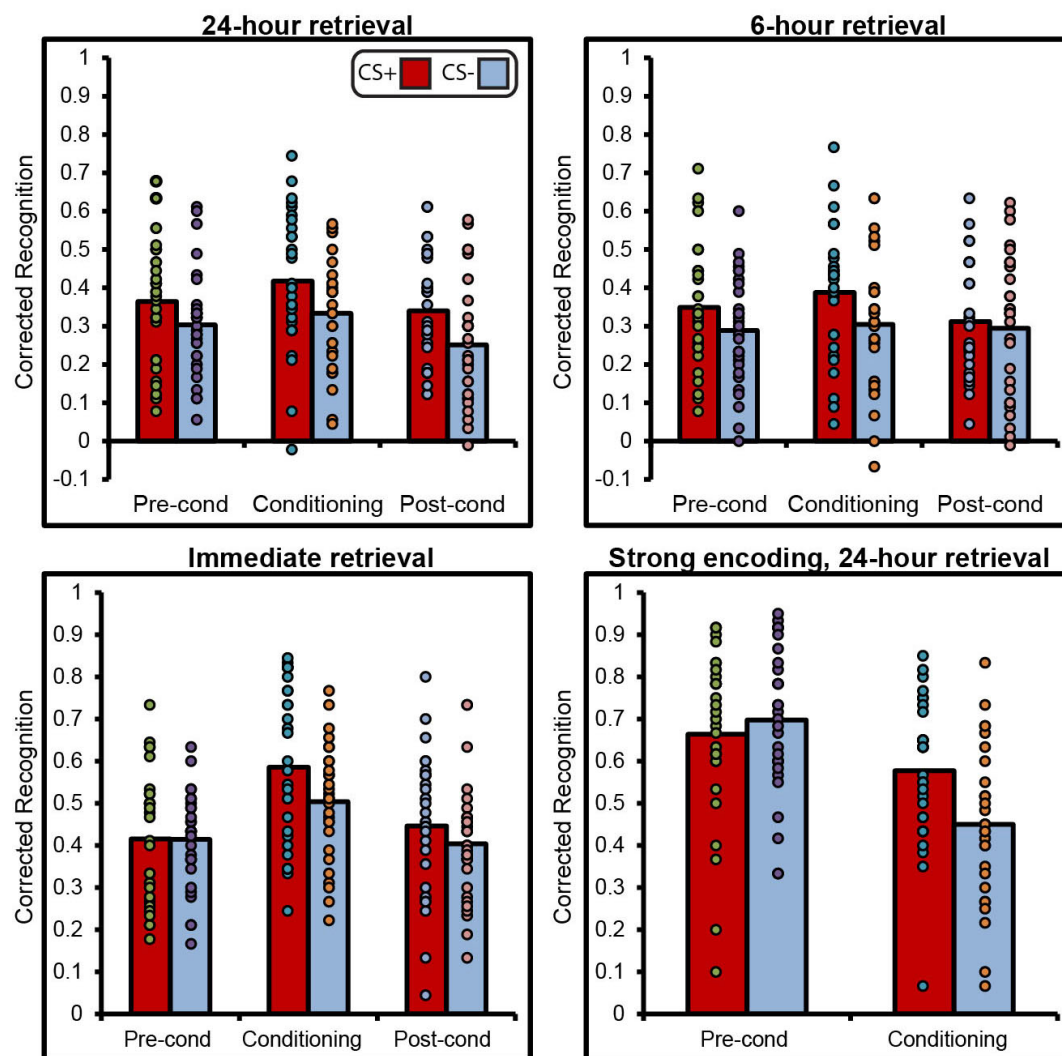
Extended Data Figure 2 | Individual data points overlaid on group means. Note that some data points overlap for subjects with the same memory score.
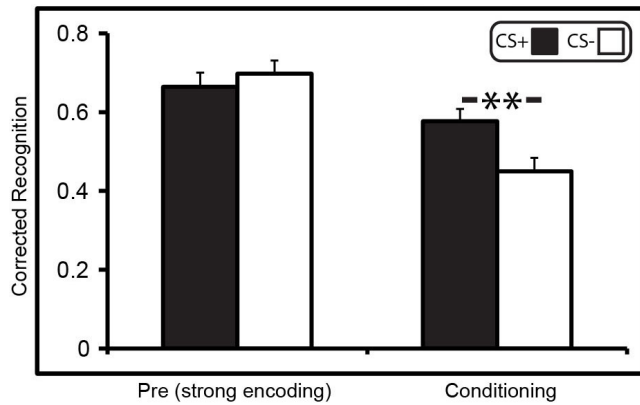
Extended Data Figure 3 | Recognition memory performance in the 24-h retrieval, strong-encoding group. Memory at 24-h retrieval for the strong-encoding group showed no difference in recognition memory between the CS^+ and CS^- items encoded before fear conditioning. Memory enhancements for stimuli encoded during fear conditioning were greater for CS^+ than CS^-, replicating the other three experiments. Error bars are s.e.m. As detailed in the main text, the $CS \times$ phase interaction was significant. $**P < 0.01$, two-tailed t-tests.

Extended Data Table 1 | 24-h retrieval memory results, mean proportion of responses

CS+	DO	MO	MN	DN	CS-	DO	MO	MN	DN
Pre-cond	0.46	0.194	0.122	0.224	Pre-cond	0.426	0.183	0.163	0.228
Conditioning	0.529	0.178	0.112	0.181	Conditioning	0.421	0.218	0.135	0.226
Post-cond	0.407	0.222	0.143	0.228	Post-cond	0.358	0.199	0.182	0.261
False Alarms	0.118	0.171	0.249	0.462	False Alarms	0.131	0.175	0.24	0.454

DO, definitely old; MO, maybe old; MN, maybe new; DN, definitely new.

Extended Data Table 2 | 6-h retrieval memory results, mean proportion of responses

CS+	DO	MO	MN	DN	CS-	DO	MO	MN	DN
Pre-cond	0.418	0.237	0.162	0.183	Pre-cond	0.374	0.222	0.178	0.226
Conditioning	0.48	0.213	0.117	0.19	Conditioning	0.387	0.223	0.168	0.222
Post-cond	0.4	0.218	0.17	0.212	Post-cond	0.363	0.223	0.192	0.221
False Alarms	0.133	0.173	0.238	0.456	False Alarms	0.134	0.169	0.248	0.449

DO, definitely old; MO, maybe old; MN, maybe new; DN, definitely new.

Extended Data Table 3 | Immediate retrieval memory results, mean proportion of responses

CS+	DO	MO	MN	DN	CS-	DO	MO	MN	DN
Pre-cond	0.435	0.217	0.166	0.182	Pre-cond	0.411	0.231	0.145	0.213
Conditioning	0.659	0.164	0.084	0.093	Conditioning	0.519	0.213	0.13	0.138
Post-cond	0.469	0.215	0.157	0.159	Post-cond	0.41	0.222	0.185	0.183
False Alarms	0.09	0.148	0.249	0.513	False Alarms	0.067	0.161	0.255	0.517

DO, definitely old; MO, maybe old; MN, maybe new; DN, definitely new.

Extended Data Table 4 | Strong-encoding, 24-h retrieval memory results, mean proportion of responses

CS+	DO	MO	MN	DN	CS-	DO	MO	MN	DN
Pre-cond	0.632	0.176	0.131	0.061	Pre-cond	0.646	0.208	0.102	0.044
Conditioning	0.496	0.228	0.168	0.109	Conditioning	0.360	0.247	0.217	0.177
False Alarms	0.043	0.104	0.309	0.544	False Alarms	0.028	0.126	0.311	0.536

DO, definitely old; MO, maybe old; MN, maybe new; DN, definitely new.