COGNITIVE NEUROSCIENCE

Stress disrupts hippocampal integration of overlapping events and memory inference in humans

Kai A. Schüren[1], Nicole L. Varga[2], Hendrik Heinbockel[1], Alison R. Preston[2,3,4], Benno Roozendaal[5,6], Lars Schwabe[1]*

Integrating related events in memory is essential for building knowledge that extends beyond direct observation and enables flexible inference. Here, we show that acute stress impairs inference by both reducing the degree to which past memories are reactivated during new learning and leading to their differentiation, rather than integration, in hippocampus. Adults learned A-B associations on day 1 and underwent a stress or control manipulation before learning overlapping B-C associations on day 2, with A-C inference tested thereafter. We demonstrate that stress reduces hippocampal reactivation of A elements during B-C learning, and lower reactivation was directly correlated with impaired A-C inference. Representational similarity analysis revealed that stress increases neural dissimilarity between overlapping A and C elements in the hippocampus, indicating pattern differentiation and a representation as discrete events. Our findings demonstrate that acute stress hampers a key memory integration mechanism, with broad implications for educational, legal, and clinical settings.

INTRODUCTION

Human memory is remarkably flexible, allowing us to maintain an up-to-date and coherent model of our environment. A key feature of this flexibility is the ability to integrate distinct experiences that share overlapping elements, forming structured knowledge representations that allow inferring novel relationships (*1*–*3*). For instance, if a friend shows you their new, light blue Vespa, and you later see the same scooter parked outside the university library, you might infer that your friend is studying inside. This ability to link related memories—a process referred to as memory integration and known to depend on the hippocampus (*4*, *5*)—is a fundamental cognitive mechanism that supports memory inference based on shared information (*6*, *7*). Beyond its fundamental relevance for cognition in general, the accuracy of memory integration has important implications for several applied contexts. In legal settings, the impaired integration of overlapping memories can lead to false inferences and wrongful accusations (*8*). In education, building cohesive memory structures by combining related information underpins conceptual understanding and predicts long-term academic outcome (*1*, *9*). Furthermore, mnemonic integration is highly relevant to mental health, as impairments in linking related experiences are characteristic of several psychiatric disorders, such as psychosis (*10*, *11*) and anxiety disorders (*12*, *13*). Thus, understanding the factors that influence our capacity to link memories is crucial.

Acute stress is a potent modulator of memory (*14*–*16*), with substantial implications for both educational and mental health outcomes (*17*–*20*). Decades of research have shown that acute stress can enhance memory consolidation while impairing memory retrieval, and both of these effects appeared to be particularly pronounced for emotionally arousing material (*21*–*24*). While most research on stress and memory has focused on consolidation and retrieval processes, initial evidence suggests that acute stress may also influence memory flexibility (*25*–*27*). Brain regions that are critical for memory integration—especially the hippocampus (*4*, *28*)—have a particularly high density of receptors for stress mediators such as glucocorticoids and norepinephrine (*29*, *30*). Through the action of these stress mediators, stress can disrupt hippocampal processing (*14*, *31*) and hence interfere with hippocampal reactivation of overlapping memory content during new learning and updating mechanisms that promote integration. However, it remains unknown whether stress affects the integration of events that share overlapping elements, allowing inferences about relationships between events that have not been experienced together, and if so, which neural mechanisms underlie such stress-induced changes in mnemonic integration.

In this preregistered study, we tested the hypothesis that acute psychosocial stress before learning of overlapping mnemonic content affects the ability to integrate memories and, consequently, memory-based inference. Using repeated functional magnetic resonance imaging (fMRI) combined with multivariate decoding and representational similarity analysis (RSA), we further sought to elucidate the neural mechanisms underlying stress effects on memory integration in the human brain. To this end, we used an associative inference task probing memory integration (*32*). In this task, healthy participants first learned A-B image associations (e.g., a friend and a light blue Vespa) on day 1. On the following day, they underwent either a psychosocial stressor [Trier Social Stress Test (TSST); (*33*)] or a control manipulation immediately before encoding B-C image associations (e.g., the same Vespa and the university library), thus testing whether stress may interfere with A element reactivation during B-C encoding. Approximately 1 hour later, we assessed participants' inference ability by testing A-C associations (e.g., friend and library). Given that stress effects on memory are typically stronger for emotionally salient material (*21*, *22*, *34*), A images were drawn from emotionally arousing (e.g., fear or threat related) or neutral categories, and emotional valence was rated by the participants at the end of the experiment. We predicted that acute stress before B-C learning would impair mnemonic integration and inference. This impairment could potentially stem from difficulties retrieving A-B and B-C pairs, which we therefore evaluated following the A-C inference test.

[1]Department of Cognitive Psychology, Universität Hamburg, Hamburg, Germany. [2]Department of Neuroscience, University of Texas at Austin, Austin, TX 78712, USA. [3]Center for Learning and Memory, University of Texas at Austin, Austin, TX 78712, USA. [4]Department of Psychology, University of Texas at Austin, Austin, TX 78712, USA. [5]Department of Medical Neuroscience, Radboud University Medical Center, Nijmegen, Netherlands. [6]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
*Corresponding author. Email: lars.schwabe@uni-hamburg.de

However, we assumed that a mere stress-induced retrieval deficit would not fully account for the predicted inference impairment. Instead, we proposed that stress would affect how new experiences are organized relative to prior, related memories. Specifically, we hypothesized that stress would disrupt reactivation of A elements during B-C learning, thereby hindering the integration of A, B, and C elements into a coherent memory trace. To test this hypothesis, A stimuli were drawn from two categories (i.e., faces and scenes) that are known to have distinct neural underpinnings (*35*, *36*). We trained a classifier on a separate task to identify these category-specific neural signatures and applied it to the B-C learning phase to quantify A element reactivation. In addition, we considered a complementary mechanism involving pattern integration or differentiation processes reflected in representational changes of individual memory elements A and C, which shared an overlapping association (B). To assess this, we presented all A and C stimuli individually on day 1 (preexposure) and again on day 2 (postexposure). Increases in representational similarity after learning would indicate memory integration, whereas decreases in representational similarity would point to memory differentiation [or repulsion; (*2*, *3*, *32*)]. This analysis of neural A and C element representations provides also important insights into how acute stress affects memory, either by increasing integration, thus leading to overgeneralization, or by decreasing it, thus reducing memory flexibility. By combining measures of representational change and memory reactivation, our experimental design extends our understanding of how stress influences the neural organization of overlapping events.

RESULTS

To unravel the impact of acute stress on memory integration of overlapping events and the underlying neural mechanisms, 121 healthy participants completed an associative inference task (*32*) within a 2-day fMRI design (Fig. 1). On day 1, participants completed a category localizer and afterward an item-level preexposure task in the MRI scanner, in which they saw all stimuli that were subsequently used as A and C elements. Thereafter, they learned 24 A-B associations as part of the associative inference task outside the scanner. On day 2, ~24 hours later, participants underwent either the TSST—a mock job interview widely regarded as a gold standard in experimental stress research (*33*)—or a control manipulation, immediately before learning 24 B-C associations that overlapped with previously learned A-B pairs and 12 nonoverlapping X-Y associations. They then completed a postexposure task, assessing learning-related representational changes of A and C elements, and a surprise A-C inference test inside the scanner. At the end of the experiment, participants rated the emotionality of the stimuli used in the task.

Successful A-B learning in the stress and the control group

During the A-B encoding task on day 1, participants were repeatedly presented 24 A-B image pairs across four runs. Immediate associative memory of these image pairs was tested in a three-alternative forced choice (3AFC) test after each of the runs. In these test trials, participants had to choose the one of three displayed B images that has previously been paired with the cued A image. Because previous research suggested that emotional material would be particularly sensitive to stress effects, we manipulated the emotionality of the A image to be neutral or negative. Participants' valence ratings confirmed that negative images were rated as significantly more negative [$F(1, 117) = 1001.55$, $P < 0.001$, $\eta^2G = 0.750$] and more arousing [$F(1, 117) = 324.58$, $P < 0.001$, $\eta^2G = 0.520$] than neutral ones. Notably, while we selected images to be emotionally neutral or negative, participants' ratings revealed that some of the images were experienced as emotionally positive, in particular neutral images (range of mean ratings across participants: 3.85 to 9.03; for negative images: 2 to 5.7). However,

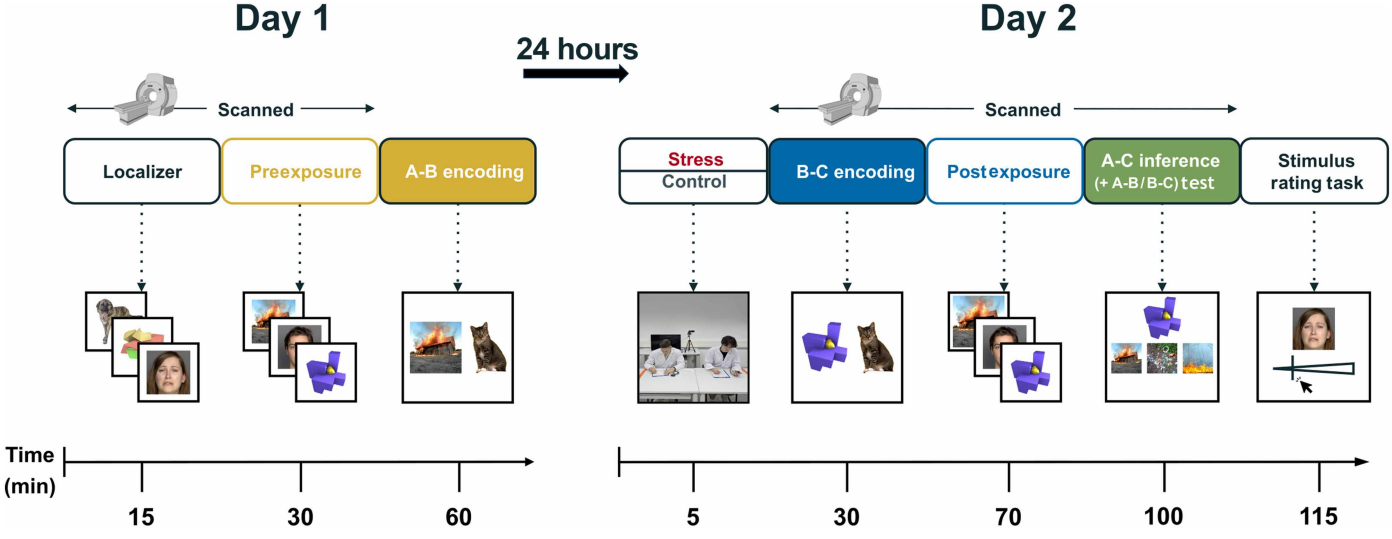

Fig. 1. Experimental procedure. Day 1 of the experiment included a category localizer task and an item-level preexposure task in which the A (negative and neutral faces) and C class stimuli (negative and neutral scenes) that would appear in the associative inference task were presented inside the MRI scanner. Afterward, participants completed the A-B encoding (24 image pairs) and test phase of the associative inference task outside of the MRI scanner. Approximately 24 hours later, day 2 started with a psychosocial stressor or a control manipulation. Afterward, participants underwent the B-C encoding phase (24 B-C pairs and 12 X-Y pairs), a postexposure task, and an A-C inference test inside the MRI scanner. Last, participants rated the valence and arousal of the images featured in the associative inference task. Face stimuli from the FACES database (*66*) with informed consent. Scene images from the Nencki Affective Picture System (NAPS) (*67*). Object stimuli generated in MATLAB as per (*32*). Animal photographs were taken by the first author. Individuals depicted in the TSST photo provided written informed consent for publication of their images. MRI graphic created in BioRender. K. Schüren (2026) https://BioRender.com/qh5mioz.

groups did not differ in emotional valence ratings [$F(1, 117) = 0.025$, $P = 0.874$, $\eta^2G < 0.001$; see figs. S1 and S2].

Figure 2A shows the memory performance for both groups, depending on the emotionality (negative versus neutral) of the involved A elements. Results suggest that both the stress and the control groups significantly improved across the four learning runs [main effect run: $F(1.55, 181.73) = 189.5$, $P < 0.001$, $\eta^2 = 0.327$] and achieved strong memory of the A-B associations at the end of the learning session. There were no main or interaction effects including the factor group, suggesting that participants that were exposed to the stressor on the subsequent day and those exposed to the control manipulation the next day learned the A-B pair associations equally well (all $F < 1.16$, all $P > 0.783$). The emotionality of the A-B pairs had also no significant effect on memory performance (main effect emotionality and emotionality × run interaction: both $F < 2.34$, both $P > 0.129$). Notably, reaction times in the A-B test trials were slightly slower in the stress group relative to the control group (see Supplementary Text).

Successful stress induction

On day 2, immediately before learning the B-C image pairs, participants were exposed to either a psychosocial stressor (TSST) or a control manipulation. Analyses of the subjective mood ratings revealed that participants' mood and calmness decreased after exposure to the TSST but not after the control manipulation [group × time interaction: good mood versus bad mood: $F(3, 228) = 3.06$, $P = 0.029$, $\eta^2 = 0.014$; calmness versus restlessness: $F(2.62, 202.08) = 8.04$, $P < 0.001$, $\eta^2 = 0.043$; and wakefulness versus tiredness: $F(2.66, 191.79) = 0.27$, $P = 0.821$, $\eta^2 = 0.002$; Fig. 3A and table S1]. In addition, participants of the stress group rated the experimental manipulation as significantly more stressful, difficult, and unpleasant than controls (all $t < -4.77$, all $P < 0.001$; table S2).

At the physiological level, systolic and diastolic blood pressure as well as heart rate increased in response to the TSST but not in response to the control manipulation (time × group interactions: all $F > 16.22$, all $P < 0.001$, all $\eta^2 > 0.039$). As shown in Fig. 3 (C to E), these autonomic measures were significantly elevated in the stress relative to the control group during the experimental manipulation [~8 min after stressor onset; all $t(115) > -4.62$, all Pcorr. < 0.001] and for systolic and diastolic blood pressure also immediately after the manipulation [~20 min after stressor onset; all $t(116) < -2.80$, all Pcorr. < 0.024]. Groups did not differ in these measures at baseline nor later after the TSST and control manipulation, respectively (~25 min after stressor onset; all $t > -0.74$, all $P > 0.356$).

Last, also salivary cortisol increased in the stress group but not in the control group [time × group interaction: $F(3.09, 333.2) = 6.32$, $P < 0.001$, $\eta^2 = 0.022$]. Figure 3F shows that there was a marked

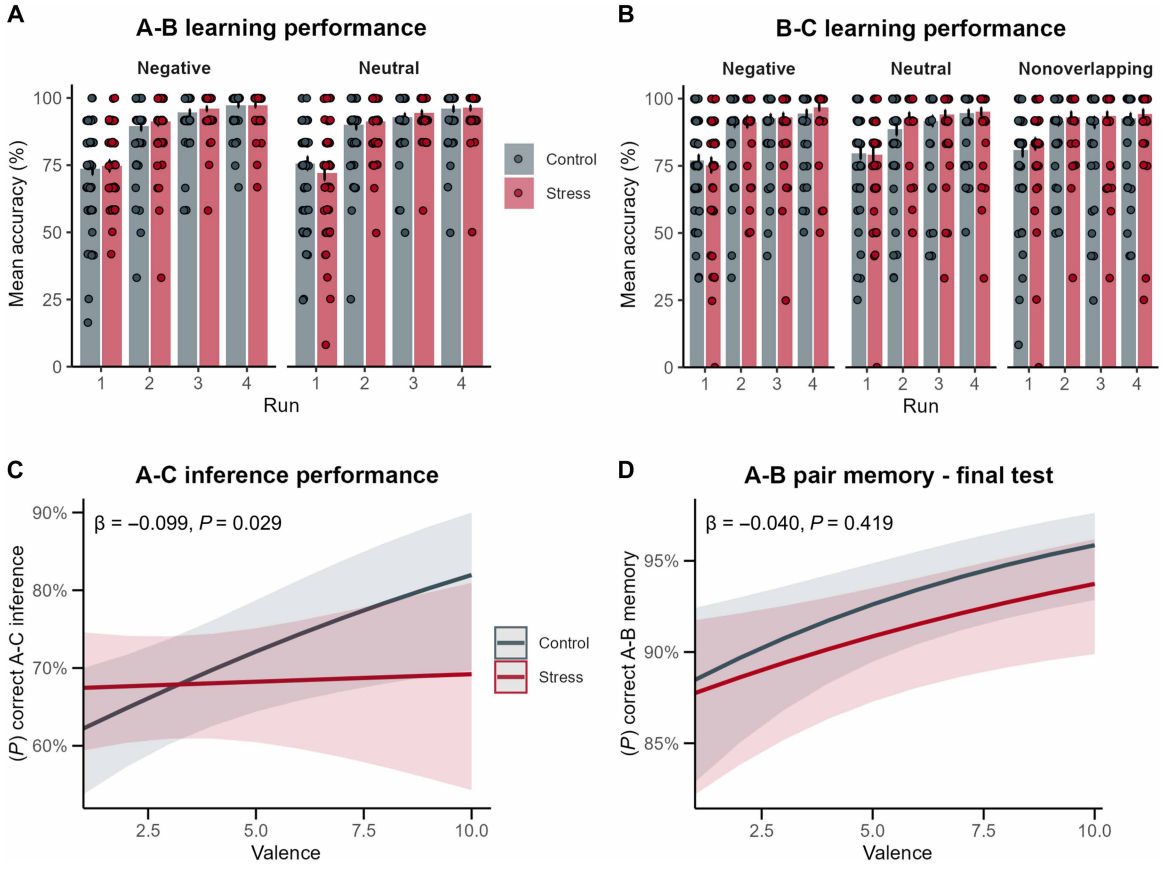

Fig. 2. Stress and perceived image valence impair associative inference. (**A**) Participants of the stress group and participants of the control group did not differ in A-B pair memory across the four learning runs on day 1 nor (**B**) in the four B-C pair learning runs on day 2. (**C**) Stress impaired inference, depending on the subjective valence of the corresponding A images. Ratings range from 1 (very negative) to 10 (very positive). (**D**) A-B memory performance on day 2 was enhanced for more negatively rated A-B pairs but not significantly affected by stress. Error bars represent SEM. The β and *P* values refer to the group × valence interaction.

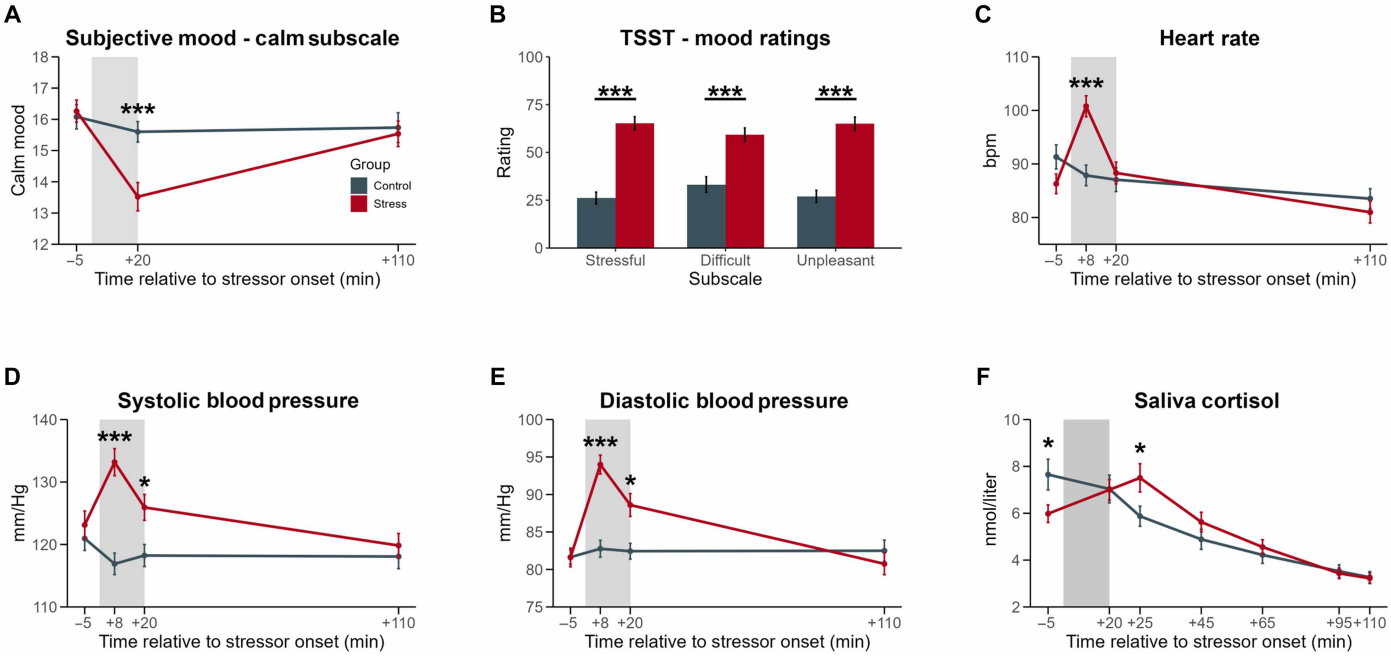

Fig. 3. Successful stress induction. (**A**) Participants in the stress group rated their mood to be less calm immediately after completion of the TSST (+20 min) compared to the control group (higher scores reflect a calmer mood). (**B**) The level of stress, difficulty, and unpleasantness experienced in the stress group during the TSST was significantly higher than in the control group. (**C**) Heart rate measured as beats per minute (bpm) increased in the stress group during the stress manipulation. (**D** and **E**) Elevated systolic and diastolic blood pressure during and after the TSST. (**F**) Increased cortisol response 25 min after the onset of the stress manipulation. Error bars represent SEM. *$P < 0.05$ and ***$P < 0.001$.

decrease in cortisol levels across the day 2 session, which is expected given the known diurnal rhythm of cortisol (*37*). However, while cortisol concentrations decreased sharply in the control group [mean baseline-to-peak increase (BPi) = −1.47 nM], there was even an increase from before to after the TSST in the stress group (mean BPi = 1.34 nM).

Stress leaves B-C learning unaffected

As for the A-B learning task, recognition memory for the B-C image pairs was tested at the end of each of the four learning runs. Performance increased significantly across runs [$F(2.23, 258.51) = 77.39$, $P < 0.001$, $\eta^2 = 0.176$], especially between run 1 and run 2 as can be seen in Fig. 2B. Both the stress and control groups learned the B-C pairs very well, without group differences [all $F < 1.81$, all $P > 0.181$, all $\eta^2 < 0.002$; see Fig. 2B], providing evidence that stress did not affect the memory of overlapping B-C pairs. Learning of new X-Y pairs was also not affected by stress [all $F < 0.74$, all $P > 0.109$, all $\eta^2 < 0.004$]. Reaction times during the B-C test trials were similar for both groups (see Supplementary Text). The absence of stress effects on B-C learning might be related to the high level of performance, which may have been facilitated by the familiarity with the associative encoding procedure or the overlapping B element.

Stress impairs inference depending on perceived A element emotionality

Participants' associative inference accuracy was assessed in a surprise A-C test approximately 1 hour after the end of the B-C encoding task. In this A-C test, participants had to infer the correct association between A and C images that were indirectly related via the corresponding B image. To test whether A-C inference was influenced by the stress manipulation and the emotionality of A images, we conducted general linear mixed model (GLMM) analyses. The model with the best fit included the predictor group (stress versus control), emotionality (negative versus neutral), subjective valence ratings on trial level, and day 2 A-B as well as B-C pair memory performance on trial level. In addition, two-way interactions of all predictors were included in the model. While the model showed no main effect of group ($\beta = 0.317$, $P = 0.249$; fig. S4), we observed a significant interaction between group and valence ($\beta = -0.099$, $P = 0.029$). As shown in Fig. 2C, inference accuracy was higher in control participants for associations that contained A elements that were subjectively experienced as emotionally positive (slope of valence for control group: $\beta = 0.103$, $P = 0.012$), in line with previous findings suggesting that positive emotion facilitates associative learning (*38*, *39*). Individuals that underwent the stress manipulation, however, showed no such modulation of inference accuracy by valence (slope of valence for the stress group: $\beta = 0.004$, $P = 0.918$) and performed particularly worse than the control group when inferences involved A elements that were rated as emotionally positive. The model included also A-B and B-C memory performance, thus the stress-related change in inference cannot be attributed to a mere impairment in memory retrieval. Arousal ratings for A class stimuli did not predict associative inference accuracy (main effect arousal and arousal × group interaction: both $P > 0.314$, both $|\beta| < 0.045$). Neither was inference accuracy significantly predicted by saliva cortisol BPi (main effect cortisol and cortisol × group interaction: both $P > 0.512$, both $|\beta| < 0.226$), suggesting that cortisol alone may not account for the observed stress effects and that other stress-related mediators likely contribute to the stress-related inference impairment. Last, reaction times in A-C inference test trials were comparable between groups (see Supplementary Text).

Beyond the test of A-C inference, we also assessed day 2 A-B and B-C retrieval performance. We ran GLMMs on A-B pair memory performance including the same predictors as for the A-C inference analysis. While a group did not significantly predict A-B pair memory performance (see fig. S3A), we observed a significant main effect of the valence ratings ($\beta = -0.123$, $P = 0.001$). As displayed in Fig. 2D, participants were more likely to recall the A-B association if the subjective valence rating for the A image was more positive. While the perceived emotionality of the A images predicted day 2 A-B pair memory performance, neither the stress manipulation nor the emotionality of the associated A images significantly affected performance in the B-C recognition trials (all $P > 0.461$; see fig. S3B).

Neural activity and connectivity during B-C encoding remains unaffected by stress

Using fMRI, we measured brain activity during B-C learning on day 2. Contrasting brain activity during the encoding of B-C image pairs with brain activity during the encoding of novel X-Y pairs revealed significant clusters in the ventromedial prefrontal cortex (vmPFC; [20, 32, 10], $t = 5.84$, $P_{\text{FWE-corr}} = 0.002$), inferior frontal gyrus ([−46, 32, 26], $t = 6.69$, $P_{\text{FWE-corr}} < 0.001$), dorsolateral PFC ([−42, 2, 56], $t = 8.94$, $P_{\text{FWE-corr}} < 0.001$), and the dorsal anterior cingulate cortex ([−4, 24, 42], $t = 7.49$, $P_{\text{FWE-corr}} < 0.001$). Increased activity in these regions during B-C encoding seems plausible as previous research has demonstrated the involvement of the hippocampus, surrounding brain areas and the mPFC in associative learning (*1*). Moreover, we analyzed seed-to-voxel functional connectivity, in particular the hippocampus, during B-C encoding trials. Contrasting B-C learning trials with X-Y learning trials showed that the functional connectivity between the hippocampus and the medial frontal cortex was increased during B-C learning ([0, 54, −2], $P_{\text{FWE-corr}} = 0.023$). This increase in functional coupling has previously been shown to be associated with increased inference accuracy, as enhanced connectivity during overlapping pair encoding is likely involved in the linking of A-B and B-C pairs (*40*). In addition, in line with our behavioral findings, stress did not affect the brain activity and functional connectivity during the encoding of overlapping B-C pairs. We obtained also no significant differences between B-C pairs overlapping with neutral versus negative A elements.

Stress impairs A element reactivation during B-C encoding

Our study aimed primarily to test the idea that stress impairs memory inference by reducing the reactivation and integration of prior experiences during the encoding of overlapping events. For this purpose, we leveraged multivariate decoding analyses assessing whether stress interferes with the A category reactivation during corresponding B-C encoding (Fig. 4A). Notably, we used A stimuli from two stimulus classes, faces and scenes, that are known to have distinct neural signatures (*41*–*43*) and trained a classifier to distinguish these categories in a separate localizer task that was presented before learning on day 1. We applied an L2-penalized logistic regression analysis, decoding scene and face images in the hippocampus, ventral temporal cortex (VTC), mPFC, parahippocampal gyrus, and perirhinal cortex. In a first step, a leave-one-run-out cross-validation was applied to determine whether classifier accuracy in each region of interest (ROI) is above chance level in the localizer task and whether the regions should be considered for the following decoding analysis. Classifier accuracy was determined by the number of correct predictions made for the random test block in relation to total amount of category predictions in the other five mini-blocks. The results of the cross-validation showed that classifier performance across all participants significantly exceeded chance level in the left perirhinal cortex [$M = 58.54$, SD = 17.64, $t(119) = 5.31$, $P < 0.001$], right perirhinal cortex [$M = 65.69$, SD = 21.12, $t(119) = 8.14$, $P < 0.001$], VTC [$M = 94.24$, SD = 10.90, $t(119) = 94.21$, $P < 0.001$], the right hippocampus [$M = 53.40$, SD = 17.57, $t(119) = 32.98$, $P < 0.001$], and the left hippocampus [M = 55.28, 17.75, $t(119) = 33.81$, $P < 0.001$]. Consequently, the decoding analysis was restricted to these brain regions, applying the trained classifier to the trial-wise data of the first B-C encoding run. Critically, groups did not differ in classifier accuracy in the localizer task in any of our ROIs [all $t(118) > -0.68$, all $P > 0.099$].

When applying the classifier to the B-C encoding task, classifier accuracy was significantly above chance level (target versus nontarget category on trial level) in the right hippocampus ($t = 121.79$, $P < 0.001$) and in the left hippocampus ($t = 112.88$, $P < 0.001$), suggesting that the linked A image categories were reactivated during B-C pair learning. In the VTC and the perirhinal cortex, however, classifier accuracy was not significantly above chance level during B-C learning. Nevertheless, both ROIs were considered for further trial-level analyses.

Next, we tested the influence of the stress manipulation on trial-level reactivation strength during B-C learning trials, operationalized as single-trial category-level classifier evidence (logits). Using linear mixed models (LMMs), the factor group, emotionality, A image category (scene versus face), as well as trial-level valence and arousal ratings and A-B and B-C (day 2) pair memory were used to predict A category reactivation. For the hippocampus, the best fitting model included group and valence ratings and their interaction effect in the model while also accounting for variance based on the image categories. This model showed a significant main effect of group, indicating significantly reduced category-level reactivation in the stress group compared to the control group ($\beta = -0.005$, $P_{\text{corr}} = 0.019$; Fig. 4B). In other words, stressed participants reactivated the corresponding A element category less in the hippocampus when encoding overlapping B-C pairs. The predictors valence ratings ($\beta = -0.0001$, $P = 0.535$), group × valence interaction ($\beta = -0.0004$, $P = 0.102$), and image category ($\beta = 0.0015$, $P = 0.052$) were not significant. Follow-up tests showed that A category reactivation was not reliably different between right and left hippocampi, speaking against a lateralized effect in the hippocampus ($\beta < -0.001$, $P = 0.940$). When the same LMMs were applied to the VTC and perirhinal cortex, none of the predictors reached significance in any of the models.

In a next step, we tested whether A category reactivation in the hippocampus successfully predicted A-C inference, thus testing the functional relevance of the A element (category) reactivation that was impaired by stress. Adding hippocampal reactivation strength to the behavioral model predicting A-C inference accuracy improved the model fit and showed a significant main effect of hippocampal reactivation on A-C inference ($\beta = 13.724$, $P = 0.041$; Fig. 4C), suggesting that lower A element category reactivation is associated with impaired inference accuracy. This association was observed across groups and did not reliably differ between groups (group × reactivation: $\beta = -12.769$, $P = 0.219$; see fig. S5), suggesting that successful reactivation of associated content during encoding of overlapping information is associated with enhanced inference accuracy, irrespective of stress. However, as reported above, acute stress significantly reduced this reactivation.

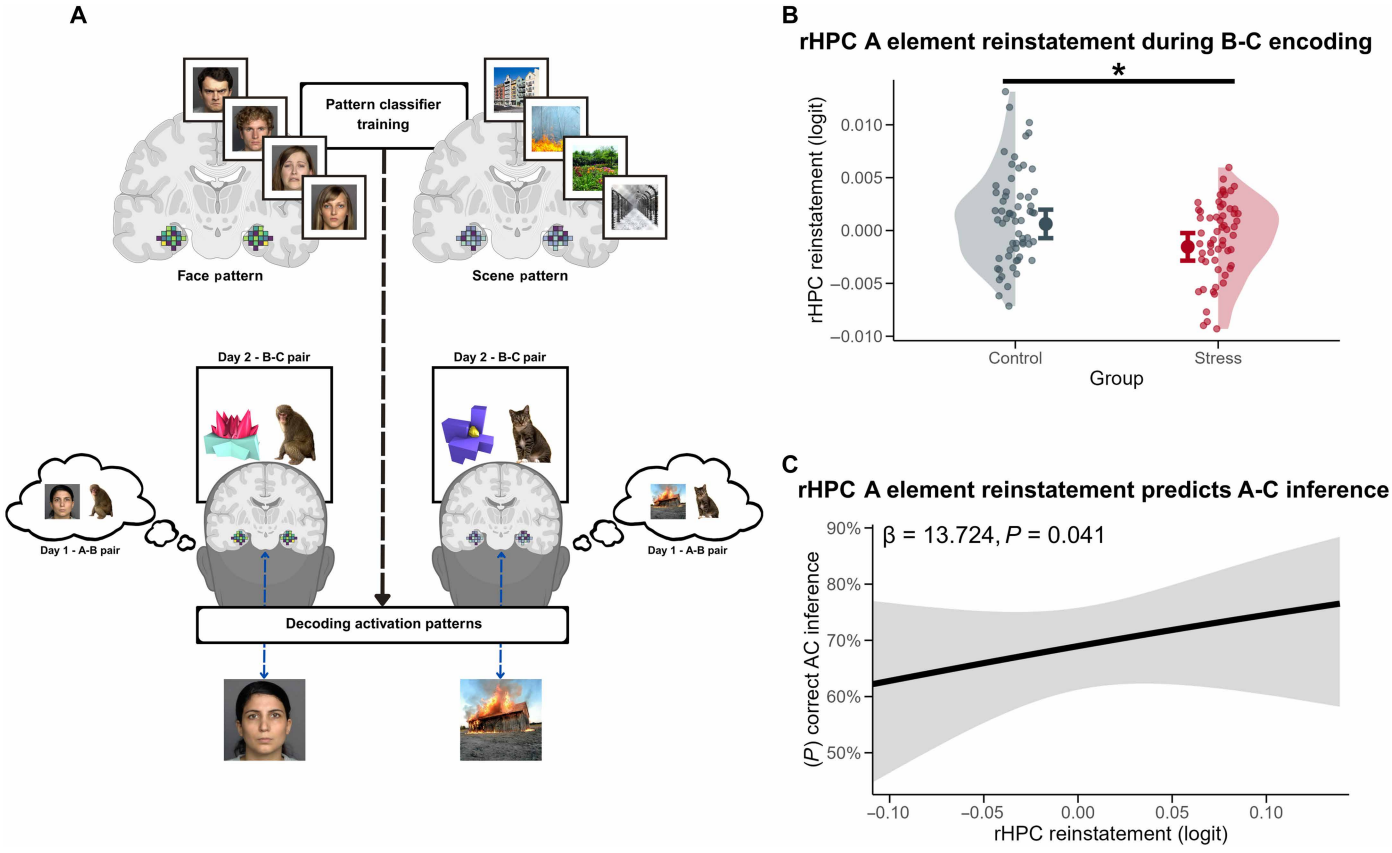

Fig. 4. Stress effect on reactivation of associated A image categories in the hippocampus during B-C trials. (**A**) Visualization of the neural decoding approach. (**B**) The reactivation of overlapping A category images during the first run of the B-C learning phase is impaired by acute stress exposure. Higher logit values indicate stronger classifier evidence for the corresponding category on each trial. (**C**) Reactivation of the A category images during the B-C learning phase significantly predicts the behavioral measure of A-C inference. Note that while (B) and (C) show the effect in the right hippocampus (rHPC) for illustrative purposes, there were no reliable differences between right and left hippocampus. Face stimuli from the FACES database (*66*) with informed consent. Scene images from NAPS (*67*). Object stimuli generated in MATLAB as per (*32*). Animal photographs were taken by the first author. Coronal brain slice and head graphics created in BioRender. K. Schüren (2026) https://BioRender.com/qh5mioz. Error bars represent SEM. *$P < 0.05$.

Stress increases neural dissimilarity of linked A-C elements in the hippocampus

To assess a potential role of memory integration and differentiation processes in the memory for overlapping events, we tested whether the neural representational similarity between A and C elements changes depending on whether these are linked through an overlapping B element and whether stress modulates such representational changes. Participants were presented these A and C elements on day 1, before A-B learning, as well as after B-C learning on day 2. We generated representational similarity matrices (RSMs) of pairwise, Fisher *z*-transformed Spearman correlation coefficients of all image presentations during the pre- and postexposure runs of the experiment for the hippocampus (Fig. 5A). Subsequently, mean similarity values for within-triad (A and C elements that were linked via an overlapping B element) and across-triad comparisons (A and C elements that were not linked) were extracted for image combinations containing negative A elements and neutral A elements. We applied LMMs to test the effects of group (stress versus control), emotionality of the A images (negative versus neutral), phase (pre- versus postexposure), and link (within-triad versus across-triad) on the representational pattern similarity of the images. These analyses revealed a significant group × emotionality interaction ($\beta = 0.002$, $P = 0.036$) showing lower pattern similarity for negative items in the stress group versus the control group ($\beta = 0.001$, $P = 0.035$; neutral: $\beta < 0.001$, $P = 0.914$) in the hippocampus. As shown in Fig. 5C, we obtained a significant interaction effect between group and link specifically in the postexposure session ($\beta = -0.003$, $P = 0.049$), indicating that the reduction in pattern similarity in the stress group was specific for within-triad comparisons. For the preexposure session, there was no such effect ($\beta = -0.001$, $P = 0.229$; Fig. 5B), nor were there any significant effects in the hippocampus (all $|\beta| < 0.002$, all $P > 0.265$). Follow-up tests showed, however, that the postexposure pattern similarity was not reliably different between the left and right hippocampus, arguing against a lateralized effect ($\beta < -0.001$, $P = 0.712$). No modulation of neural pattern similarity was found for analyses of the mPFC and perirhinal cortex. Note that we also analyzed the pre-to-post difference for the between versus within-triad difference between stress and control groups. This interaction, however, was not significant ($\beta < -0.001$, $P = 0.740$), presumably due to insufficient power for this high-level interaction. Additional searchlight analyses within ROIs did not reveal any effects of group or A element emotionality on pre-to-post similarity change either.

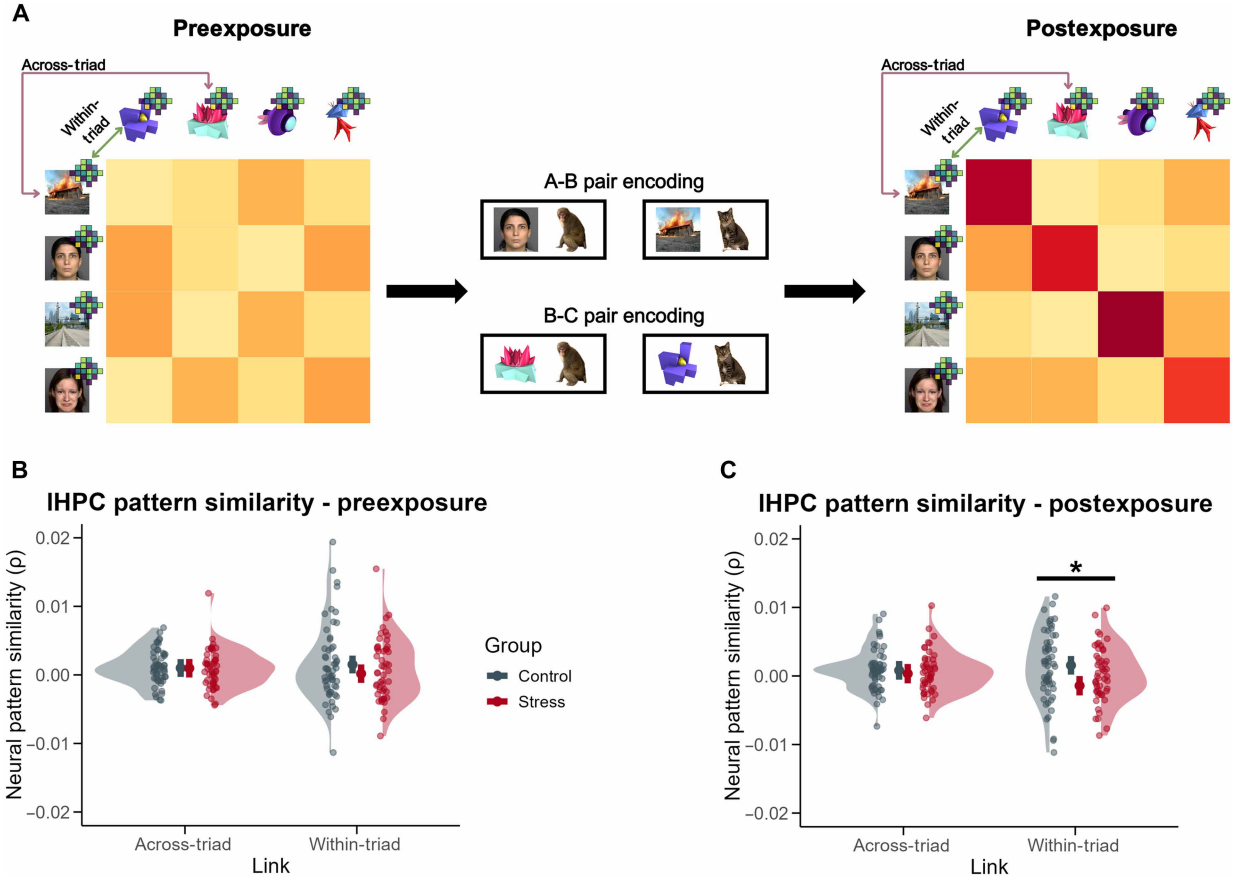

Fig. 5. Representational similarity in the left hippocampus. (**A**) Depiction of the analytical approach. (**B**) Before A-B and B-C learning, neural similarity between individual A and C elements was comparable between groups and within- versus across-triad pairs. (**C**) After A-B and B-C learning, the neural similarity decreased specifically for within-triad A and C elements in the stress group, suggesting stress-induced pattern differentiation. Note that while (B) and (C) show the effect in the left hippocampus (lHPC) for illustrative purposes, there were no reliable differences between right and left hippocampus. Face stimuli from the FACES database (*66*) with informed consent. Scene images from NAPS (*67*). Object stimuli generated in MATLAB as per (*32*). Animal photographs were taken by the first author. Error bars represent SEM. *$P < 0.05$.

Last, adding A category reactivation measures of the hippocampus to the postexposure model showed no significant effect of the reactivation on pattern similarity ($\beta = -0.046$, $P = 0.471$), suggesting that the changes in A element reactivation during overlapping B-C learning and the representational change of individual A and C elements sharing an overlapping B element are uncorrelated and likely reflect distinct mechanisms relevant for memory integration. We did not identify a modulation of the behavioral A-C inference scores by similarity changes in any of the ROIs.

Increased hippocampus, amygdala, and mPFC activity during associative inference

In addition, we analyzed brain areas involved in mnemonic inference. To this end, we contrasted brain activity in inference (i.e., A-C) and memory retrieval (i.e., A-B and B-C) trials. Moreover, we explored whether inference-related univariate activity was influenced by stress or A element emotionality. For the overall contrast of inference minus retrieval, our ROI analyses revealed significant clusters in the bilateral hippocampi (left: [−20, −14, −22], $t = 6.59$, $P_{\text{FWE-corr}} < 0.001$; right: [22, −10, −10], $t = 9.94$, $P_{\text{FWE-corr}} < 0.001$), bilateral amygdalae (left: [−20, −6, −14], $t = 6.68$, $P_{\text{FWE-corr}} < 0.001$; right: [20, −8, −18], $t = 8.20$, $P_{\text{FWE-corr}} < 0.001$), mPFC ([0, 48, −18], $t = 7.07$, $P_{\text{FWE-corr}} < 0.001$), and the parahippocampal cortex ([24, −34, −18], $t = 14.72$, $P_{\text{FWE-corr}} < 0.001$), suggesting that these areas are more strongly activated during inference compared to retrieval trials. In addition, a whole-brain analysis revealed significant clusters in the occipital lobe, inferior frontal lobe, orbitofrontal cortex, and fusiform gyrus (see table S3). The stress manipulation did not significantly influence this inference-related neural activity, suggesting that the observed inference impairment is mainly driven by the mnemonic integration during the encoding of overlapping events rather than changes in neural activity during inference. An additional analysis focusing exclusively on correct associative inference trials revealed largely overlapping activation clusters for the A-C versus A-B/B-C pair memory contrast. No additional effects of group or emotionality were observed in either the whole-brain or ROI analyses.

DISCUSSION

The ability to integrate separate memories that share overlapping elements is essential for memory inference, with important implications in legal, educational, and clinical contexts. In this study, we investigated whether acute stress influences this fundamental memory

process and, if so, which neural mechanisms underlie stress-induced changes in inference. We found that stress before learning events containing elements overlapping with existing memories impaired participants' ability to associate related episodes, leading to deficits in associative inference—particularly for stimuli rated as emotionally positive. This stress-induced impairment in associative inference was accompanied by reduced neural reactivation of existing memories in the hippocampus during the encoding of overlapping content. In addition, we observed a stress-induced increase in pattern differentiation for overlapping memories in the hippocampus. Together, these findings demonstrate that acute stress impairs the neural integration of overlapping events and hence inference, possibly due to a diminished capacity to reinstate overlapping memories during integrative hippocampal encoding.

Previous research on stress and memory has primarily focused on the consolidation and retrieval of individual memories, showing that acute stress can enhance memory formation (*21*, *22*) but impair retrieval (*23*, *24*). Our findings extend this literature significantly by demonstrating that stress before learning of overlapping events also disrupts integrative encoding, which is important for associative inference. The inference deficit in the stress group, compared to the control group, was specific to stimuli that participants subjectively rated as emotionally positive, consistent with prior evidence that emotional memory content is particularly susceptible to stress effects (*16*, *21*, *22*). This modulation by stimulus emotionality has been attributed to interactions between glucocorticoids and noradrenergic activity in the brain (*34*). In particular, these stress hormones modulate basolateral amygdala functioning and the interplay with the hippocampus and neocortical brain regions, thereby mediating stress induced memory impairments. Previous studies have shown that positive emotional memory content can enhance associative learning by expanding attentional focus (*38*, *39*). While we selected emotionally negative and neutral stimuli, valence ratings of the images included a positive dimension, and participant subjective valence ratings showed that several of the neutral stimuli were experienced as rather positive. In light of these findings, it might well be that the greater associative inference accuracy for more positively rated items observed in the control group could be the result of positive emotional modulation and that acute stress interfered specifically with this modulation. However, further studies including explicitly both emotionally positive and negative stimuli are needed to investigate this relationship as well as potential differences between positive and negative emotionally arousing events.

Our results further align with studies showing that acute stress impairs memory flexibility, as reflected, for instance, in reduced schema-based learning or decreased spatial navigation adaptability (*26*, *27*, *44*). Together, these findings suggest that stress effects on memory are broad and not limited to enhanced memory formation and impaired memory retrieval. Crucially, changes in memory formation or retrieval alone cannot explain the present results. Stress did not impair the acquisition of B-C pairs and left brain activity and functional connectivity during B-C learning unaffected, indicating that the inference deficit was not due to a failure to encode the overlapping content. Similarly, stress did not impair retrieval of the A-B or B-C associations in the final test, possibly because these were highly overlearned. Including retrieval performance as a covariate in our model did not eliminate the stress effect on inference. Thus, the observed inference deficit cannot be attributed solely to encoding or retrieval impairments and goes beyond the previously established effects of stress on memory. Notably, the absence of a stress effect on A-B memory suggests that potential context shifts between A-B learning (outside the scanner), and subsequent A-C and A-B testing (inside the scanner) did not substantially influence performance or hippocampal engagement.

Rather than reflecting deficits in encoding or retrieval alone, our fMRI decoding results point to a stress-induced reduction in memory integration, through which related memories are represented in overlapping neural code. Specifically, we show that stress disrupted the reactivation of the corresponding A element category in the hippocampus during the learning of the overlapping B-C image pairs, a first step critical for memory integration (*28*). The hippocampus is known to play a key role in associative inference (*4*). In particular, previous research showed that integrative encoding of overlapping events in the hippocampus is key to generalization across experiences (*45*). Reactivation of overlapping memory content is thought to underlie memory integration and, consequently, successful inference (*46*). Consistent with this view, our results showed that stronger hippocampal A category reactivation during learning of overlapping events predicted successful associative inference. The stress-induced reduction in memory reactivation aligns with previous findings that stress hippocampal neuroplasticity and function are sensitive to stress (*31*, *47*). Notably, stress has been shown to promote a shift from flexible hippocampal memory toward more rigid, cue-dependent striatal memory (*14*, *17*), which may hinder the relational processing and integration of overlapping experiences. We therefore propose that stress impairs associative inference by disrupting spontaneous retrieval of prior, related memories during new encoding, which has been shown to influence the degree to which related events become integrated within the same representational network (*32*).

In line with this proposal, our RSA points to a potential complementary mechanism underlying the effects of stress on mnemonic integration. Specifically, we found that stress resulted in increased neural dissimilarity in the neural representation of A and C elements that shared an overlapping B element. This effect was specific to associations involving an emotionally negative event. Focusing the analysis on the postexposure phase then showed a decrease in pattern similarity in the stress group particularly for linked A-C associations. The observed increased neural dissimilarity indicates pattern differentiation and may emerge when episodes are represented as separate, nonoverlapping events (*48*). Thus, these stress-induced representational changes provide further evidence for reduced integration of overlapping elements after stress. Again, this stress effect was observed within the hippocampus, implying that under stress, this region not only fails to sufficiently reactivate overlapping content but may actively differentiate related episodes. Notably, increased hippocampal pattern differentiation after acute stress has been observed in spatial navigation tasks, where it contributes to more precise spatial codes and facilitates navigational accuracy (*49*). This enhancement of pattern differentiation, however, could come at the expense of pattern integration capacities required to form an integrated map of the spatial environment. The reactivation of the A category during B-C learning and the increase in neural dissimilarity between the A and C elements were uncorrelated, suggesting that these may reflect independent mechanisms—one potentially affecting overlapping content reactivation and the other governing the degree of differentiation among overlapping representations. It is to be noted, however, that this stress-induced representational effect did not reach statistical significance when controlling for the preexposure representation (i.e., representational change), presumably due to a lack of statistical

power. Moreover, we did not observe a direct relationship between the neural pattern similarity change in the hippocampus and our behavioral measure of associative inference, potentially because this behavioral measure lacks the sensitivity to detect fine-grained neural modulations.

While our findings suggest that stress primarily affected the hippocampal reactivation and integration of A category during the encoding of overlapping events (i.e., B-C), it is important to note that the timing of our stress manipulation was specifically chosen to target this process. Accordingly, cortisol was elevated during B-C encoding but had returned to baseline by the time of the A-C inference test. Therefore, our results do not preclude the possibility that stress may also influence processes relevant at the time of inference—such as the reactivation of the overlapping B element—when stress occurs immediately before inference testing. Moreover, it remains unclear whether memory integration and its modulation by stress require explicit knowledge of the link between events. Future studies could include a test of participants' awareness of this link.

In conclusion, our findings demonstrate that acute stress can impair the ability to integrate separately encoded but overlapping events, leading to deficits in associative inference. This impairment is likely driven by a stress-induced disruption of memory reactivation in the hippocampus during the encoding of overlapping information, resulting in unintegrated representations of related experiences. Given the ubiquity of stressful events in everyday life, such deficits in memory integration and inference have relevant implications across many domains, most importantly mental disorders. Notably, while impaired integration of overlapping events may hinder flexible inference, increased memory differentiation may also serve adaptive functions, such as reducing interference. Our findings suggest that under stress, the brain prioritizes the distinct representation of individual episodes over the formation of connected knowledge structures.

MATERIALS AND METHODS

Preregistration

This study, including its hypotheses and analyses, was preregistered before the end of data collection at the Open Science Framework (https://doi.org/10.17605/OSF.IO/4Y96V).

Participants and experimental design

One hundred twenty-one healthy, right-handed volunteers participated in this experiment. Volunteers fluent in German were screened in a standardized interview for exclusion criteria, including a lifetime history of any neurological or psychiatric disorders, thyroid, renal, or cardiovascular conditions, any contraindications for MRI measurements, consumption of nicotine, prescription medications, or illicit drugs, and, for women, pregnancy, lactation, or the use of hormonal contraceptives. An a priori power analysis (*50*) (*G**Power, 3.1.9) indicated that 120 participants would be sufficient to detect a medium-sized effect [$d = 0.55$; based on (*25*)] of the two-level group factor (stress versus control) on associative inference, with a power of 0.90. All participants provided written informed consent before participating in the study and received monetary compensation upon completion. The experimental procedure was approved by the local ethics committee of the Faculty of Psychology and Human Movement Science at the University of Hamburg (2020_343_Schwabe).

In a mixed design with the two-level between-subjects factor group (stress versus control), the two-level within-subject factor emotionality (negative versus neutral A stimuli in the associative inference task), and the two-level within-subject factor associative link (within versus between triad), participants were pseudo-randomly assigned to either the stress group ($n = 60$; 30 female; mean age = 23.8, SD = 4.1) or the control group ($n = 61$; 31 female; mean age = 25.1, SD = 4.4), ensuring a comparable number of men and women in each group.

To control for any potential differences in anxiety, depression, and chronic stress symptoms between the two experimental groups that participants were assigned to, we applied the Trier Inventory of Chronic Stress (*51*), the Community Assessment of Psychic Experiences–positive dimension (*52*), the Childhood Trauma Questionnaire–short form (*53*), and the Inventory of Depression and Anxiety Symptoms-I (*54*) before the start of the experiment. Independent sample *t* tests that we ran on all subscales of the questionnaires showed that participants assigned to the stress group and participants assigned to the control group did not differ in any of the symptoms (tables S4 to S7).

Experimental procedure and tasks

Day 1: Localizer task

To train a classifier to distinguish between the neural signatures of the stimulus categories used in the associative learning task, participants first completed a category localizer consisting of three runs. The localizer task included the block-wise presentation of four different image categories: scenes (2/3 manmade, 1/3 natural), faces (1/2 male, 1/2 female), artificial three-dimensional objects, and animals, with 48 images presented per category. In addition, the faces were either emotionally neutral or negative (sad, fearful, or angry facial expressions), and the scenes were either emotionally neutral (peaceful/preserved urban or natural scenes) or negative (natural disasters, destruction, or pollution scenes). We selected scenes that evoked negative or neutral emotion while excluding depictions of humans or large objects, as such features could interfere with neural pattern decoding analyses. Images used for the localizer task were randomly drawn from four category-specific image pools for each participant and did not subsequently reappear in any of the other tasks of the experiment. Each run of the localizer involved the presentation of eight mini-blocks (two per image category), with each mini-block consisting of nine images from one category. Images were presented for 2.5 s, and a fixation cross was shown for 0.5 s between images (jittered between 0.3 and 0.8 s). To ensure adequate attention during the localizer task, participants were instructed to press a button whenever one of the images was presented twice in a row (1-back memory task). One of the nine images in each mini-block was always a duplicate of one of the other eight images.

Day 1: Preexposure task

Following the localizer task, participants underwent a preexposure phase inside the MRI, during which they were presented with images that would later reappear in the associative inference task. Specifically, participants viewed 24 images from the subsequent A stimulus class, featuring faces and scenes, as well as 24 images from the subsequent C stimulus class, featuring objects. As in the localizer task, the images in the A stimulus class were divided into negative and neutral images. The preexposure task consisted of four runs, with each image shown twice per run in a pseudo-random order, ensuring that items from the same A-B-C triad would always have a minimum of two intervening items in between. The timing and trial orders of each run were kept identical between the pre- and postexposure phases to control for the influence of carry-over effects on the neural activity

measures of temporally adjacent items. Following the same rationale, the pseudo-random order of stimulus presentation ensured that same triad stimuli were always intertwined by two items from different triads. Images were presented for 1 s, with an intertrial interval (ITI) of 2.5 to 3.5 s. To ensure participants' attention during the preexposure task, the fixation cross during the ITI changed color from black to either green or blue, and participants were instructed to indicate the correct color by pressing a button.

Day 1: A-B pair learning

As part of the associative inference task, participants first learned a series of A-B pairs outside the MRI. In this A-B learning task, participants were instructed to memorize 24 pairs of A class (faces or scenes) and B class (animals) stimuli. Unbeknownst to participants, A-B image pairs would later be linked to B-C image pairs, and an unannounced A-C inference test would be applied toward the end of the experiment. During the study phase, the A and B images were presented next to each other, with the A image displayed on the left side of the screen and the B image on the right. Each pair was shown for 3.5 s, followed by a 0.5-s fixation cross. Immediately following the study phase, participants completed a brief test phase, in which they had to select the correct image pair in a 3AFC test. In this test phase, each of the previously shown A images appeared once as a cue at the top of the screen, with three B images displayed beneath it, including the one which was previously paired with the cue. Test trials were also counterbalanced to ensure that the two foils displayed alongside the correct pair would always be from the same stimulus category and emotionality (e.g., face/neutral). The decision time for test trials was 5 s, and participants received visual feedback on the correctness of their choice. Each run of the A-B learning consisted of both study and test phases. Participants completed four runs in total, with each run including the same 24 A-B pairs.

Day 2: Stress manipulation

At the beginning of day 2, ~24 hours after the end of day 1 (M = 24.09; SD = 0.82; no significant difference between groups), participants underwent either the TSST (*33*) or a control manipulation. The TSST is a well-established psychosocial stressor in which participants are exposed to a mock job interview, consisting of a free speech and a mental arithmetic task. During a 3-min anticipation phase, participants were instructed to prepare a free speech about why they are the ideal candidate for a job tailored to their interests. After the 5-min free speech, participants were asked to perform a 5-min mental arithmetic task, in which they had to count backward from 2043 in steps of 17. Throughout the free speech and the mental arithmetic task, participants were evaluated by a panel of two experimenters, dressed in white laboratory coats, who were introduced as experts in behavioral analysis. The panel remained cold and nonreinforcing. In addition, participants were videotaped and could see themselves on a screen placed behind the panel.

In the control condition, participants gave a speech about a topic of their choice (e.g., their last vacation or a book they read) and performed a simple mental arithmetic task (counting in steps of 15). There was no panel during the control manipulation, and participants were not videotaped.

To assess the success of the stress manipulation, saliva samples were taken at multiple time points across the two experimental days applying the passive drool method using Salivettes (Sarstedt, Germany) to measure the stress hormone cortisol (*55*). An initial baseline sample was taken at the beginning of day 1, before any of the experimental tasks. Similarly, a baseline sample was obtained on day 2, before the stress or control manipulation. Further, saliva samples were taken immediately after the stress or control manipulation, 5 min later while preparing participants for the MRI, and after the second run of the B-C learning task. Three additional saliva samples were taken: before the postexposure task, before the final A-C test, and after the completion of the MRI session on day 2. Salivary cortisol concentrations (nanomoles per liter) were analyzed via antibody detection using a luminescence immunoassay (IBL; Germany). All testing took place in the morning, with sessions starting between 7:25 a.m. and 11:30 a.m. to control for the diurnal rhythm of cortisol.

In addition to the saliva samples, participants' heart rate and blood pressure were measured at five different time points: baseline measures at the beginning of day 1, at the beginning of day 2, 8 min after the onset of the stress and control manipulation, immediately after the manipulation (+20 min after TSST onset), and after the MRI session (+110 min after TSST onset). Subjective mood was assessed using a German multidimensional mood scale (*56*) (MDBF) at the beginning of day 1, the beginning of day 2, immediately after the TSST, and at end of the scanning session (+110 min after TSST onset). In addition, participants were asked to rate, on a scale from 0 ("not at all") to 100 ("very much"), how stressful, difficult, and unpleasant they had experienced the TSST and control manipulation, respectively. Baseline measures of salivary cortisol, heart rate, blood pressure, and the MDBF from day 1 were analyzed using independent sample t tests to control for potential group differences before the experimental manipulation (see Supplementary Text). Moreover, we ran mixed analyses of variance (ANOVAs) including the factors time and group to check whether groups differed in response to the experimental manipulation across the course of day 2 of the study.

Day 2: B-C pair learning

Immediately after the end of the stress or control manipulation, participants completed the B-C learning task inside the MRI scanner. The procedure of this B-C learning task was essentially identical to the previous A-B learning task, except for different stimulus timing due to the MRI measurement. The 24 B class images that were previously associated with A images during the A-B learning task on day 1 were now paired with 24 C class images (objects). In addition, 12 novel X-Y image pairs, containing B class (i.e., animal) images that had not been presented before, were introduced. These novel X-Y pairs served as a baseline for the effect of stress on new encoding, behaviorally and for univariate and functional connectivity analyses of fMRI data. As in the A-B learning task, the 24 B-C pairs and the 12 X-Y pairs were presented across four runs. C class images were shown on the left of the screen, while B class images were displayed on the right. Each B-C image pair was presented for 3.5 s, with an ITI of 8.5 s during which a fixation cross was shown. Each run consisted of a study phase, during which all B-C pairs were presented in a randomized order, followed by a 3AFC test phase to probe associative memory for the learned B-C pairs. For the B-C test trials, C images were shown at the top of the screen, and the three B image options were displayed beneath. As for the A-B test phase, test trials were counterbalanced by stimulus category and additionally by pair type (B-C and X-Y pairs). Participants had 5 s to respond but did not receive feedback about the correctness of their choice.

Day 2: Postexposure task

Immediately after the B-C learning task, participants underwent a postexposure phase. Here, the 24 A and 24 C images shown during the associative inference task (i.e., A-B and B-C learning, respectively)

were presented again across four runs, following the same procedure as during the preexposure task on day 1.

Day 2: A-C inference test

At the end of the fMRI scanning session on day 2, ~95 min after the onset of the stress or control manipulation, participants completed a surprise A-C inference test, which probed the indirect A and C image association established through the overlapping B image. This A-C test was announced as a "free association test," in which participants should select the person (represented by the face stimulus) who would most likely own the cued object or the scene where the cued object could most likely be found. During the A-C inference test, C images were shown at the top of the screen, and three A image options were displayed underneath, with one of the options being the correct stimulus, indirectly associated via the overlapping B stimulus. In addition to A-C inference trials, this final test also probed memory for the 24 A-B pairs learned on day 1 and for the 24 B-C pairs learned earlier on day 2. The A-C test consisted of 72 trials in total, including 24 inference trials, 24 A-B pair memory trials, and 24 B-C pair memory trials. To ensure that inference accuracy was not influenced by prior exposure to the respective A-B and B-C image pairs, four A-C trials were presented before the eight corresponding A-B and B-C memory trials. The test trials for the A-B and B-C pairs followed the same procedure as the test trials during the A-B and B-C learning tasks on day 1 and day 2, respectively. For each of the three trial types, participants had 10 s to respond.

Day 2: Stimulus emotionality rating

At the end of day 2, after leaving the MRI, participants completed a 10-min stimulus rating task. In this task, participants were instructed to rate the valence and arousal of the face and scene images that were used as A class stimuli in the experiment. The task consisted of 120 trials, presenting the 60 face images and the 60 scene images that participants had seen in the localizer and associative inference task. For each image, participants rated the valence on a 10-point Likert scale from 1 (very negative) to 10 (very positive) and the arousal on a scale from 1 (very calm) to 10 (very arousing). Although we selected A class stimuli to be emotionally neutral versus negative, we still used a valence scale from "very negative" to "very positive" because we considered the possibility that the subjective perception of the participants could deviate from these categories. Especially the neutral stimuli could be potentially perceived as rather positive, in particular when contrasted with the emotionally negative stimuli. Participants' ratings showed that some of the A class stimuli were perceived as emotionally positive (see figs. S1 and S2).

Behavioral data analysis

To assess the success of the stress manipulation, salivary cortisol, heart rate, blood pressure, and mood data were subjected to mixed-design ANOVAs with the between-subjects factor group (stress versus control) and the within-subject factor time point of measurement. Performance in the four A-B and B-C learning runs (treating B-C and X-C test trials separately) was also analyzed by applying ANOVAs including the group factor, as well as the within-subject variables run and emotionality (negative versus neutral; for B-C learning runs, emotionality was based on the associated A elements). Associative inference accuracy was analyzed using GLMMs. Using a stepwise approach, the predictor group, emotionality (negative versus neutral image), and stress responsiveness, measured as BPi in salivary cortisol, were added to the models. In addition, subjective valence and arousal ratings of the images in the associative inference task, as well as A-B and B-C pair memory from the final test session, were added to the regression models. Last, A-B and B-C memory in the final test was analyzed using GLMMs, including group, emotionality, and valence and arousal ratings as predictors.

fMRI data acquisition and preprocessing

MRI data were acquired with a 3T Siemens Magnetom Prisma scanner using a 64-channel head coil. For functional imaging, we applied T2*-weighted echo-planar imaging (EPI) sequences (TR = 2000 ms, TE = 30 ms) with a voxel size of 2 mm^2 and a slice thickness of 2 mm. In addition to the functional volumes, structural images were obtained at the beginning of day 1, including a T1-weighted image using a magnetization-prepared rapid gradient echo (MPRAGE) sequence (TR = 1900 ms, TE = 2.34 ms) and high-resolution T2-weighted volume (TR = 13,150 ms, TE = 82 ms) of the hippocampus.

Functional and structural MRI data were preprocessed using standard steps, as implemented in SPM12 (statistical parametric mapping; www.fil.ion.ucl.ac.uk/spm/), including spatial realignment, slice time correction, coregistration to the structural images, segmentation, normalization to the Montreal Neurological Institute (MNI) standard space, and smoothing with a 6-mm Gaussian kernel.

ROI analyses

ROI masks were derived primarily from the Harvard-Oxford cortical and subcortical atlas (thresholded at 50%), unless specified otherwise. Preregistered ROIs for analyses of task-evoked activation included the bilateral (i.e., left and right) hippocampus, amygdala, mPFC, and the parahippocampal and perirhinal cortices, as previous research indicated these brain regions to be involved in associative memory and inference (*1*, *57*, *58*). For functional connectivity analyses, the bilateral hippocampus, amygdala, and mPFC were used as seed regions.

For the RSA and multivariate pattern analysis (MVPA), which were conducted in native space, ROI masks were back-transformed using the inverse deformation fields obtained from the segmentation preprocessing step. ROIs for these analyses included the bilateral hippocampus, mPFC, and parahippocampal and perirhinal cortices. In addition, a VTC mask was included exclusively for the MVPA and was derived as described in (*59*, *60*). The perirhinal cortex mask was derived from the probabilistic medial temporal lobe atlas (*61*) (https://identifiers.org/neurovault.collection:3731).

Univariate fMRI data analysis

To assess neural activity associated with inference (versus retrieval) in the A-C test, we derived general linear models (GLMs) for each of the participants using SPM12. This analysis included the trial types of the A-C test (inference versus associative memory retrieval trials) and the emotionality of the stimuli that were tested (negative versus neutral) as regressors. Moreover, item-level valence ratings were added as a parametric modulator to regressors in the first-level GLMs, generating modulated regressors of the conditions estimating interaction effects. At the first level, we focused on the contrast A-C trials (i.e., inference) minus A-B/B-C trials (i.e., memory retrieval) and the reverse contrast. At the second (group) level, we contrasted the stress versus control groups and analyzed potential interaction effects of the treatment with stimulus emotionality. In addition, we conducted an analysis focusing exclusively on correctly remembered A-C trials applying the same first and second-level contrasts. Note that this deviates from our preregistered approach where we indicated to contrast

correct and incorrect A-C trials instead. We chose this approach to focus specifically on neural activity associated with successful inference. Contrasting correct and incorrect trials would have mixed distinct cognitive processes, reducing interpretability and sensitivity. We performed explorative whole-brain analyses as well as analyses in predefined (and preregistered) ROIs. Univariate brain activity during B-C encoding trials was analyzed in a similar fashion, contrasting overlapping B-C trials with nonoverlapping X-Y trials and contrasting B-C trials associated with negative images with B-C trials that were associated with neutral images. Last, the stress and control groups were contrasted using the aforementioned first-level contrasts. All univariate fMRI analyses were Bonferroni-corrected for the number of ROIs with a (cluster-level) significance threshold of $P < 0.05$ (FWE corrected).

Functional connectivity analysis

We performed functional connectivity analyses for the B-C learning runs of the associative inference task as well as the A-C inference (including the A-B and B-C pair recognition trials) at the end of the experiment. We conducted seed-to-voxel analyses using the CONN toolbox (*62*) implemented in MATLAB. Denoising was applied including SPM-derived motion parameters as well as physiological noise extracted from white matter and cerebrospinal fluid. Moreover, a temporal high-pass filter of 0.008 Hz was applied to reduce low-frequency drift while preserving task relevant high-frequency signals. In a first-level analysis, correlation maps of each seed ROI and every voxel in the brain were computed for each participant. Second-level analyses assessing task-related variance in connectivity were conducted using GLMs. For analyses focusing on the B-C learning phase, second-level contrasts included negative versus neutral B-C trials (based on the emotionality of the corresponding A-class images). X-C trials (B and C image pairs that did not overlap with previously learned A-B image pairs) were also contrasted with B-C trials. Last, we contrasted the stress and control groups and assessed the interaction effects of the emotionality and group. Analyses of the A-C inference test included negative versus neutral trials as a contrast, A-C trials versus A-B/B-C pair recognition trials, as well as stress versus control group contrasts. In addition, we assessed interaction effects of the aforementioned variables. The resulting statistical maps were thresholded at $P < 0.001$ with a cluster false discovery rate correction at $P < 0.05$.

Multivariate decoding analysis

To assess the neural reactivation of A class stimuli during B-C learning, we leveraged MVPA-based decoding analysis. These analyses were based on participants' realigned, slice-time corrected, native-space images. In a first step, block-wise beta estimates for the localizer task were estimated together with trial-wise betas for all B-C study phases in within-subject GLMs. The MVPA models were then derived using modified functions from the Decoding Toolbox (TDT) (*63*). To evaluate classification performance within the localizer task, a leave-one-out cross-validation was applied, training L2-penalized logistic regression models on five of the category-specific mini blocks (scenes versus faces) of the task and testing the classifier on a random sixth block. Preregistered ROIs where classifier accuracy significantly surpassed chance level (50%) were then selected for further analysis. Note that this approach deviates from our preregistered plan to identify ROIs for further analysis by comparing classifier evidence for overlapping B-C pairs with nonoverlapping X-Y pairs. The latter approach would require the assignment of random category labels to the novel, nonoverlapping pairs (which were, per definition, not paired with an A element before). Assigning arbitrary labels to these trials would render classifier accuracy computed largely uninterpretable. We therefore decided against our preregistered plan and instead included ROIs for which classifier accuracy was above chance level.

Subsequently, the pattern classifier trained on localizer data within the respective ROIs was applied to the B-C learning phase data to detect (re-)activation patterns related to the two A stimulus categories. We assessed trial-wise category reactivation by using logits, which represent the signed

In the final step, trial-wise reactivation of A image categories (i.e., faces or scenes) during overlapping B-C learning was assessed. For each trial, we extracted the classifier's probabilistic output for the face and scene categories, where 1 reflects perfect classifier evidence for the target category and 0 reflects no evidence. These probabilities were converted into logits applying log-odds transformations $\{\text{logit}(p) = \ln[p/(1-p)]\}$, yielding continuous measures of category-level evidence (i.e., face logits and scene logits). Higher positive values reflect stronger evidence for the target category. This transformation was applied to control for nonnormality following established protocols [e.g., (*64*, *65*)]. These trial-wise measures of category reactivation were then analyzed to determine whether stress and stimulus emotionality affect A stimulus reactivation during B-C learning. Note that, deviating from the preregistration, we used LMM regressions to analyze trial-wise reactivation, as we applied a continuous outcome variable (logits), rendering LMMs the more appropriate approach. The models included group and emotionality as factors, as well as the interaction between group and valence, and as covariates the baseline-to-peak cortisol levels, stimulus type, trial-wise valence and arousal ratings, and A-B and B-C pair memory performance. The decoding analysis focused on the first run of the B-C learning task because potential A element reactivation in subsequent runs may be compromised by B-C pair familiarity.

Representational similarity analysis

To assess associative learning–related changes in stimulus representations and potential stress effects on these representational changes, we leveraged RSA. Specifically, we analyzed whether the representational similarity between A and C stimuli linked by an overlapping B element would increase from pre- to postexposure, indicating potential integration processes. Conversely, we examined whether the representational similarity between A and C stimuli not linked by an overlapping B element would decrease from pre- to postexposure, indicating potential pattern differentiation (*32*). Trial-wise regressors were derived at the stimulus level by collapsing across all presentations of each image per run. The resulting onsets were then applied to GLMs run on native space images for all participants. Using TDT functions, we then generated similarity matrices using pairwise, Fisher *z*-transformed Spearman correlation indices of A and C image category activation patterns across runs, within the following preregistered ROIs: hippocampus, mPFC, perirhinal cortex, and parahippocampal cortex. Within each ROI, we then extracted within-triad (A and C elements that were linked in the associative inference task) and across-triad (A and C elements that were not linked) pattern similarity values from the resulting similarity matrices, however, exclusively for across run comparisons. We then examined the effects of stress and emotionality of the A elements, as well as the effect of linking (within versus across triad comparisons) applying LMMs. In

addition, we looked at pre- and postexposure phases individually applying LMMs that included the same factors. Last, we analyzed similarity change across phases by subtracting preexposure similarity from postexposure similarity and subtracting across-triad similarity from within-triad similarity (*32*).

To assess local representational change within the preregistered ROIs, we applied a spherical searchlight approach (radius = 3 voxels). Within each participant's native functional space, a searchlight sphere was iteratively centered on every voxel of the ROI masks. For each sphere, we extracted trial-wise activation patterns from the pre- and postlearning phases and computed all pairwise Pearson correlations between item-level patterns. Correlation coefficients were Fisher *z*-transformed and assembled into RSMs separately for each sphere.

From each RSM, we quantified within- and across-triad similarities for items associated through prior learning (A-C pairs from the same or different triads, respectively). To ensure statistical independence of activation patterns, all similarity estimates were restricted to comparisons between runs within the same phase. For each item pair, a change-in-similarity (Δ-similarity) score was computed as post- minus prelearning similarity. We then obtained, per sphere, the mean Δ-similarity for within- and across-triad comparisons, separately for negative and neutral items.

The principal contrast of interest reflected the interaction between the linking structure (within versus across triads) and the emotional category (negative versus neutral). In additional analyses, we also computed a main linking contrast that collapsed across emotional categories, quantifying the overall difference between within- and across-triad Δ-similarities regardless of valence.

To obtain voxel-wise significance maps, we implemented subject-level nonparametric permutation testing. Within each searchlight sphere, within/across labels were randomly shuffled (preserving the number of pairs per condition) for 1000 iterations, and the contrast statistic was recomputed to generate a null distribution. The observed contrast value was compared against this distribution to obtain a two-tailed *P* value, which was subsequently converted to a *z*-statistic. The resulting subject-level *S*, *p*, and *z* maps were saved in native space and then normalized to MNI space using each participant's deformation field derived from anatomical preprocessing. Group-level inference on normalized *z*-maps was performed using voxel-wise one-sample *t* tests within SPM, and significant clusters were identified using voxel-wise thresholding ($P < 0.01$ uncorrected) and cluster-extent correction ($P < 0.05$, family-wise error corrected).

Supplementary Materials

This PDF file includes:

Supplementary Text
Tables S1 to S7
Figs. S1 to S5

REFERENCES

1. M. L. Schlichting, A. R. Preston, Memory integration: Neural mechanisms and implications for behavior. *Curr. Opin. Behav. Sci.* **1**, 1–8 (2015).
2. I. K. Brunec, J. Robin, R. K. Olsen, M. Moscovitch, M. D. Barense, Integration and differentiation of hippocampal memory traces. *Neurosci. Biobehav. Rev.* **118**, 196–208 (2020).
3. A. F. de Sousa, A. Chowdhury, A. J. Silva, Dimensions and mechanisms of memory organization. *Neuron* **109**, 2649–2662 (2021).
4. N. W. Morton, K. R. Sherrill, A. R. Preston, Memory integration constructs maps of space, time, and concepts. *Curr. Opin. Behav. Sci.* **17**, 161–168 (2017).
5. H. C. Barron, H. M. Reeve, R. S. Koolschijn, P. V. Perestenko, A. Shpektor, H. Nili, R. Rothaermel, N. Campo-Urriza, J. X. O'Reilly, D. M. Bannerman, T. E. J. Behrens, D. Dupret, Neuronal computation underlying inferential reasoning in humans and mice. *Cell* **183**, 228–243.e21 (2020).
6. L. Raviv, G. Lupyan, S. C. Green, How variability shapes learning and generalization. *Trends Cogn. Sci.* **26**, 462–483 (2022).
7. N. Varga, N. Morton, A. Preston, Schema, inference, and memory. OSF [Preprint] (2022). https://doi.org/10.31234/osf.io/m9adb.
8. T. D. Albright, Why eyewitnesses fail. *Proc. Natl. Acad. Sci.* **114**, 7758–7764 (2017).
9. N. L. Varga, A. G. Esposito, P. J. Bauer, Cognitive correlates of memory integration across development: Explaining variability in an educationally relevant phenomenon. *J. Exp. Psychol. Gen.* **148**, 739–762 (2019).
10. K. Armstrong, S. Avery, J. U. Blackford, N. Woodward, S. Heckers, Impaired associative inference in the early stage of psychosis. *Schizophr. Res.* **202**, 86–90 (2018).
11. A. Zadbood, Y. Tang, W. Su, H. Hu, G. Capichioni, S. Yang, J. Wang, V. P. Murty, C. Gasser, O. Bein, L. Hui, Q. Jia, T. Zhang, Y. Hong, M. F. Green, J. Wang, D. C. Goff, L. Davachi, Impaired hippocampal circuitry and memory dysfunction in schizophrenia. *Nat. Mental Health* **3**, 332–345 (2025).
12. S. Lissek, A. N. Kaczkurkin, S. Rabin, M. Geraci, D. S. Pine, C. Grillon, Generalized anxiety disorder is associated with overgeneralization of classically conditioned fear. *Biol. Psychiatry* **75**, 909–915 (2014).
13. O. Laufer, D. Israeli, R. Paz, Behavioral and neural mechanisms of overgeneralization in anxiety. *Curr. Biol.* **26**, 713–722 (2016).
14. L. Schwabe, E. J. Hermans, M. Joëls, B. Roozendaal, Mechanisms of memory under stress. *Neuron* **110**, 1450–1467 (2022).
15. C. Sandi, M. T. Pinelo-Nava, Stress and memory: Behavioral effects and neurobiological mechanisms. *Neural Plast.* **2007**, 078970 (2007).
16. M. Joëls, G. Fernandez, B. Roozendaal, Stress and emotional memory: A matter of timing. *Trends Cogn. Sci.* **15**, 280–288 (2011).
17. S. Vogel, L. Schwabe, Learning and memory under stress: Implications for the classroom. *npj Sci. Learn.* **1**, 16011 (2016).
18. D. de Quervain, L. Schwabe, B. Roozendaal, Stress, glucocorticoids and memory: Implications for treating fear-related disorders. *Nat. Rev. Neurosci.* **18**, 7–19 (2017).
19. R. K. Pitman, A. M. Rasmusson, K. C. Koenen, L. M. Shin, S. P. Orr, M. W. Gilbertson, M. R. Milad, I. Liberzon, Biological studies of post-traumatic stress disorder. *Nat. Rev. Neurosci.* **13**, 769–787 (2012).
20. L. Schwabe, Memory under stress: From adaptation to disorder. *Biol. Psychiatry* **97**, 339–348 (2025).
21. B. Roozendaal, S. Okuda, D. J.-F. de Quervain, J. L. McGaugh, Glucocorticoids interact with emotion-induced noradrenergic activation in influencing different memory functions. *Neuroscience* **138**, 901–910 (2006).
22. L. Cahill, L. Gorski, K. Le, Enhanced human memory consolidation with post-learning stress: Interaction with the degree of arousal at encoding. *Learn. Mem.* **10**, 270–274 (2003).
23. D. J. de Quervain, B. Roozendaal, J. L. McGaugh, Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature* **394**, 787–790 (1998).
24. D. J.-F. de Quervain, B. Roozendaal, R. M. Nitsch, J. L. McGaugh, C. Hock, Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nat. Neurosci.* **3**, 313–314 (2000).
25. L. C. Dandolo, L. Schwabe, Stress-induced cortisol hampers memory generalization. *Learn. Mem.* **23**, 679–683 (2016).
26. T. I. Brown, S. A. Gagnon, A. D. Wagner, Stress disrupts human hippocampal-prefrontal function during prospective spatial navigation and hinders flexible behavior. *Curr. Biol.* **30**, 1821–1833.e8 (2020).
27. C. W. E. M. Quaedflieg, L. Schwabe, Memory dynamics under stress. *Memory* **26**, 364–376 (2018).
28. D. Zeithamova, A. L. Dominick, A. R. Preston, Hippocampal and Ventral Medial Prefrontal Activation during Retrieval-Mediated Learning Supports Novel Inference. *Neuron* **75**, 168–179 (2012).
29. B. S. McEwen, J. M. Weiss, L. S. Schwartz, Selective retention of corticosterone by limbic structures in rat brain. *Nature* **220**, 911–912 (1968).
30. J. M. H. M. Reul, E. R. de Kloet, Two receptor systems for corticosterone in rat brain: Microdistribution and differential occupation. *Endocrinology* **117**, 2505–2511 (1985).
31. J. J. Kim, D. M. Diamond, The stressed hippocampus, synaptic plasticity and lost memories. *Nat. Rev. Neurosci.* **3**, 453–462 (2002).
32. R. J. Molitor, K. R. Sherrill, N. W. Morton, A. A. Miller, A. R. Preston, Memory reactivation during learning simultaneously promotes dentate gyrus/$CA_{2,3}$ pattern differentiation and CA_1 memory integration. *J. Neurosci.* **41**, 726–738 (2021).
33. C. Kirschbaum, K.-M. Pirke, D. H. Hellhammer, The 'Trier Social Stress Test'—A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* **28**, 76–81 (1993).
34. B. Roozendaal, B. S. McEwen, S. Chattarji, Stress, memory and the amygdala. *Nat. Rev. Neurosci.* **10**, 423–433 (2009).

35. N. Kanwisher, Neural events and perceptual awareness. *Cognition* **79**, 89–113 (2001).
36. F. Tong, M. S. Pratte, Decoding Patterns of Human Brain Activity. *Annu. Rev. Psychol.* **63**, 483–509 (2012).
37. N. O. Dmitrieva, D. M. Almeida, J. Dmitrieva, E. Loken, C. F. Pieper, A day-centered approach to modeling cortisol: Diurnal cortisol profiles and their associations among U.S. adults. *Psychoneuroendocrinology* **38**, 2354–2365 (2013).
38. C. R. Madan, S. M. E. Scott, E. A. Kensinger, Positive emotion enhances association-memory. *Emotion* **19**, 733–740 (2019).
39. R. Pan, C. Gao, X. Zhu, B. Li, X. Jia, Positive emotion enhances memory by promoting memory reinstatement across repeated learning. *J. Neurosci.* **45**, e0009252025 (2025).
40. M. L. Schlichting, A. R. Preston, Hippocampal–medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neurobiol. Learn. Mem.* **134**, 91–106 (2016).
41. N. Kanwisher, G. Yovel, The fusiform face area: A cortical region specialized for the perception of faces. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **361**, 2109–2128 (2006).
42. R. Epstein, A. Harris, D. Stanley, N. Kanwisher, The parahippocampal place area: Recognition, navigation, or encoding? *Neuron* **23**, 115–125 (1999).
43. C. McCormick, E. A. Maguire, The distinct and overlapping brain networks supporting semantic and spatial constructive scene processing. *Neuropsychologia* **158**, 107912 (2021).
44. L. M. Kluen, A. Agorastos, K. Wiedemann, L. Schwabe, Noradrenergic stimulation impairs memory generalization in women. *J. Cogn. Neurosci.* **29**, 1279–1291 (2017).
45. D. Shohamy, A. D. Wagner, Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron* **60**, 378–389 (2008).
46. N. W. Morton, E. L. Zippi, A. R. Preston, Memory reactivation and suppression modulate integration of the semantic features of related memories in hippocampus. *Cereb. Cortex* **33**, 9020–9037 (2023).
47. E. J. Kim, B. Pellman, J. J. Kim, Stress effects on the hippocampus: A critical review. *Learn. Mem.* **22**, 411–416 (2015).
48. M. L. Schlichting, J. A. Mumford, A. R. Preston, Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* **6**, 8151 (2015).
49. A. Tomar, T. J. McHugh, The impact of stress on the hippocampal spatial code. *Trends Neurosci.* **45**, 120–132 (2022).
50. F. Faul, E. Erdfelder, A.-G. Lang, A. Buchner, G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
51. P. Schulz, W. Schlotz, Trierer Inventar zur Erfassung von chronischem Sre (TICS): Skalenkonstruktion, teststatistische Überprüfung und Validierung der Skala Arbeitsüberlastung. [The Trier Inventory for the Assessment of Chronic Stress (TICS). Scale construction, statistical testing, and validation of the scale work overload.]. *Diagnostica* **45**, 8–19 (1999).
52. N. C. Stefanis, M. Hanssen, N. K. Smirnis, D. A. Avramopoulos, I. K. Evdokimidis, C. N. Stefanis, H. Verdoux, J. V. Os, Evidence that three dimensions of psychosis have a distribution in the general population. *Psychol. Med.* **32**, 347–358 (2002).
53. D. P. Bernstein, J. A. Stein, M. D. Newcomb, E. Walker, D. Pogge, T. Ahluvalia, J. Stokes, L. Handelsman, M. Medrano, D. Desmond, W. Zule, Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse Negl.* **27**, 169–190 (2003).
54. D. Watson, M. W. O'Hara, K. Naragon-Gainey, E. Koffel, M. Chmielewski, R. Kotov, S. M. Stasik, C. J. Ruggero, Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment* **19**, 399–420 (2012).
55. D. H. Hellhammer, S. Wüst, B. M. Kudielka, Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* **34**, 163–171 (2009).
56. R. Steyer, P. Schwenkmezger, P. Notz, M. Eid, Testtheoretische Analysen des Mehrdimensionalen Befindlichkeitsfragebogen (MDBF). [Theoretical analysis of a multidimensional mood questionnaire (MDBF).]. *Diagnostica* **40**, 320–328 (1994).
57. G. R. I. Barker, E. C. Warburton, Object-in-place associative recognition memory depends on glutamate receptor neurotransmission within two defined hippocampal-cortical circuits: A critical role for AMPA and NMDA receptors in the hippocampus, perirhinal, and prefrontal cortices. *Cereb. Cortex* **25**, 472–481 (2015).
58. C. R. Madan, E. Fujiwara, J. B. Caplan, T. Sommer, Emotional arousal impairs association-memory: Roles of amygdala and hippocampus. *Neuroimage* **156**, 14–28 (2017).
59. H. Heinbockel, A. D. Wagner, L. Schwabe, Post-retrieval stress impairs subsequent memory depending on hippocampal memory trace reinstatement during reactivation. *Sci. Adv.* **10**, eadm7504 (2024).
60. H. Heinbockel, G. Leicht, A. D. Wagner, L. Schwabe, Post-retrieval noradrenergic activation impairs subsequent memory depending on cortico-hippocampal reactivation. *eLife* **13**, RP100525 (2025).
61. M. Ritchey, M. E. Montchal, A. P. Yonelinas, C. Ranganath, Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *eLife* **4**, e05025 (2015).
62. S. Whitfield-Gabrieli, A. Nieto-Castanon, Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* **2**, 125–141 (2012).
63. M. N. Hebart, K. Görgen, J.-D. Haynes, The Decoding Toolbox (TDT): A versatile software package for multivariate analyses of functional imaging data. *Front. Neuroinform.* **8**, 88 (2015).
64. F. R. Richter, A. J. H. Chanales, B. A. Kuhl, Predicting the integration of overlapping memories by decoding mnemonic processing states during learning. *Neuroimage* **124**, 323–335 (2016).
65. T. I. Brown, S. A. Gagnon, A. D. Wagner, Stress disrupts human hippocampal-prefrontal function during prospective spatial navigation and hinders flexible behavior. *Curr. Biol.* **30**, 1821–1833.e8 (2020).
66. C. A. C. Holland, N. C. Ebner, T. Lin, G. R. Samanez-Larkin, Emotion identification across adulthood using the Dynamic FACES database of emotional expressions in younger, middle aged, and older adults. *Cognit. Emot.* **33**, 245–257 (2019).
67. A. Marchewka, Ł. Żurawski, K. Jednoróg, A. Grabowska, The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav. Res. Methods* **46**, 596–610 (2014).

Acknowledgments: We gratefully acknowledge the assistance of C. Hiller with the programming of the task and M. B. Alcaide, A. Tiefert, A. Turlach, N. Rieck, J. Weber, L. Wedmann, E. L. Jetter, J. Cromme, B. Schleheck, and V. Zöbelin during data collection.
Funding: This study is part of the Research Training Group 2753 "Emotional Learning and Memory," funded by the German Research Foundation (DFG; grant 449640848 to L.S.).
Author contributions: Conceptualization: L.S., B.R., A.R.P., and K.A.S. Methodology: L.S., A.R.P., B.R., N.L.V., K.A.S., and H.H. Software: K.A.S. Validation: L.S., H.H., and K.A.S. Formal analysis: K.A.S. Investigation: K.A.S. Resources: L.S. and N.L.V. Writing—original draft: K.A.S. and L.S. Writing—review and editing: L.S., K.A.S., H.H., B.R., N.L.V., and A.R.P. Visualization: K.A.S. Supervision: L.S., N.L.V., and B.R. Project administration: L.S. Funding acquisition: L.S. and B.R.
Competing interests: The authors declare that they have no competing interests. **Data, code, and materials availability:** All data and code needed to evaluate and reproduce the results in the paper are present in the paper and/or the Supplementary Materials, have been permanently deposited in a publicly accessible research repository, and are available at http://doi.org/10.25592/uhhfdm.18245. This study did not generate new materials.

Submitted 11 July 2025
Accepted 13 April 2026
Published 22 May 2026
10.1126/sciadv.aea5496

Science Advances

Stress disrupts hippocampal integration of overlapping events and memory inference in humans

Kai A. Schüren, Nicole L. Varga, Hendrik Heinbockel, Alison R. Preston, Benno Roozendaal, and Lars Schwabe

Sci. Adv. **12** (21), eaea5496. DOI: 10.1126/sciadv.aea5496